

University of Vermont

ScholarWorks @ UVM

---

Graduate College Dissertations and Theses

Dissertations and Theses

---

2020

## Quantifying the health benefits of nature contact in cities across the US

Aaron J. Schwartz  
*University of Vermont*

Follow this and additional works at: <https://scholarworks.uvm.edu/graddis>



Part of the [Place and Environment Commons](#)

---

### Recommended Citation

Schwartz, Aaron J., "Quantifying the health benefits of nature contact in cities across the US" (2020).  
*Graduate College Dissertations and Theses*. 1268.  
<https://scholarworks.uvm.edu/graddis/1268>

This Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks @ UVM. It has been accepted for inclusion in Graduate College Dissertations and Theses by an authorized administrator of ScholarWorks @ UVM. For more information, please contact [donna.omalley@uvm.edu](mailto:donna.omalley@uvm.edu).

# QUANTIFYING THE HEALTH BENEFITS OF NATURE CONTACT IN CITIES ACROSS THE US

A Dissertation Presented

by

Aaron J. Schwartz

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements  
for the Degree of Doctor of Philosophy  
Specializing in Natural Resources

August, 2020

Defense Date: May 14, 2020  
Dissertation Examination Committee:

Taylor Ricketts, Ph.D., Advisor  
Chris Danforth, Ph.D., Chairperson  
Brendan Fisher, Ph.D.  
Donna Rizzo, Ph.D.  
Cynthia J. Forehand, Ph.D., Dean of Graduate College

# Abstract

Urbanization, the rise of sedentary lifestyles, and increasing screen time have led to a significant decline in nature contact, or how much time people spend in greenspace. At the same time, urban populations are experiencing declining physical and mental well-being. While nature contact has been shown to have a variety of health benefits, these benefits have not been well-quantified or verified across different geographic contexts. In addition, there is a lack of clarity around how the benefits of nature contact vary temporally (e.g. seasonally) and between different types of greenspaces. In this dissertation, I investigate the health benefits of urban parks at three spatial scales. In Chapter One, I used the Hedonometer, a sentiment analysis tool, to show that people write happier words on Twitter when visiting parks in San Francisco. The sentiment benefit, or change in average word happiness from baseline, was highest for the larger and greener regional parks. In Chapter Two, I applied similar methods to the 25 largest cities in the US and again found higher sentiment during park visits. The sentiment benefit was highest for the largest parks and during weekends and the summer. In Chapter Three, using national health surveys and a database of municipal park systems, I found that increased park access is associated with lower prevalence of negative health outcomes (including obesity and mental health) across the 500 largest cities in the US. The findings of this research support the importance of protecting and enhancing urban nature and can be used by urban planners and public health officials to better inform nature contact recommendations for growing urban populations.

# Citations

Material from this dissertation has been published in the following form:

Schwartz, A. J., Dodds, P. S., O’Neil-Dunne, J. P., Danforth, C. M., & Ricketts, T. H.. (2019). Visitors to urban greenspace have higher sentiment and lower negativity on twitter. *People and Nature*, 1(4), 476–485.

Material from this dissertation has been submitted for publication:

Schwartz, A. J., Dodds, P. S., O’Neil-Dunne, J. P., Ricketts, T. H., & Danforth, C. M.. (2020). Exploring twitter sentiment in urban parks across 25 cities in the usa. *Science Advances* (*in submission*).

## Acknowledgements

I am so grateful for the family, friends, colleagues, and mentors who have supported me the past 5 years. I would not have made it here without your guidance, collaboration, and friendship. I would also like to acknowledge the support of the National Science Foundation Graduate Research Fellowship Program and the Medina Fund.

The Gund Institute has been a fantastic home base. Thank you to the staff who have made it an amazing place to work. Extra thanks to those I've shared an office with: Joe, Hilary, Adrian, and Eva. Long live the Dork Playground.

The members the Ricketts Lab have helped this dissertation get to the finish line in more ways than one. I am lucky to have crossed paths with so many excellent scholars and people: Insu Koh, Diego Herrera Garcia, Ranaivo Rasolofson, Leif Richardson, Nitin Singh, Laura Sonter, Keri Watson, Charlie Nicholson, Jesse Gourevitch, Natalia Aristizabal, Aura Alonso-Rodriguez, Luz De Wit, Tim Treuer, Caitlin Littlefield, Maya Moore, and Eva Kinnebrew. Additional gratitude to other friends and colleagues who provided feedback on this dissertation: Lindsay Barbieri, Meg Egler, Katie Horner, Caitlin Morgan, and Kristin Raub.

I was fortunate to cross campus regularly and be part of The Computational Story Lab, an impressive group of scientists tackling challenging problems with grit, technical chops, humor, and a surprising amount of physical fitness.

My committee has provided invaluable guidance and support along the way. Thank you Donna Rizzo and Brendan Fisher for your insights. I also want to thank Peter Dodds, Jarlath O'Neil-Dunne, Matthew Browning, and Alessandro Rigolon for their contributions to this research.

I was incredibly lucky to have two fantastic advisors, Taylor Ricketts and Chris

Danforth. Thank you both for being incredible mentors and role models. I am continually impressed by the leadership, inclusivity, and sense of joy you both bring to this community daily.

The friends I found here in Burlington ensured I had fun, got outside, and danced along this journey. I am so grateful for all of you.

Grandma Marilyn, thank you for your unfailing belief in me. Ian, you are an incredible brother and my best friend. Mom and Dad, thank you for all of the support and love over the years.

# Table of Contents

Citations . . . . .	ii
Acknowledgements . . . . .	iii
List of Figures . . . . .	vii
List of Tables . . . . .	viii
<b>1 Introduction</b>	<b>1</b>
1.1 Problem Statement . . . . .	1
1.2 Background and Gaps . . . . .	2
1.3 Research Scope . . . . .	5
<b>2 Visitors to urban greenspace have higher sentiment and lower negativity on Twitter</b>	<b>7</b>
2.1 Abstract . . . . .	7
2.2 Introduction . . . . .	8
2.3 Methods . . . . .	11
2.3.1 Study Site & Data Collection . . . . .	11
2.3.2 Tweet Binning . . . . .	12
2.3.3 Sentiment Analysis . . . . .	14
2.3.4 Estimating Sentiment . . . . .	15
2.3.5 Duration Calculation . . . . .	16
2.3.6 Park Classifications and NDVI . . . . .	16
2.4 Results . . . . .	18
2.5 Discussion . . . . .	20
2.6 References . . . . .	27
<b>3 Gauging the happiness benefit of US urban parks through Twitter</b>	<b>33</b>
3.1 Abstract . . . . .	33
3.2 Introduction . . . . .	33
3.3 Methods . . . . .	36
3.3.1 Data Collection & Processing . . . . .	36
3.3.2 Sentiment Analysis . . . . .	37
3.3.3 Park Analysis . . . . .	40
3.3.4 Temporal Analysis . . . . .	40
3.4 Results . . . . .	41
3.4.1 Sentiment Analysis . . . . .	41
3.4.2 Park Analysis . . . . .	43
3.4.3 Temporal Analysis . . . . .	47
3.5 Discussion . . . . .	48
3.5.1 Sentiment Analysis . . . . .	48
3.5.2 Park Analysis . . . . .	49
3.5.3 Temporal Analysis . . . . .	50

3.5.4	Future Directions . . . . .	50
3.6	References . . . . .	51
<b>4</b>	<b>The health benefits of urban parks vary regionally across cities</b>	<b>56</b>
4.1	Abstract . . . . .	56
4.2	Introduction . . . . .	56
4.3	Methods . . . . .	59
4.3.1	Health Outcomes . . . . .	59
4.3.2	Park Access Estimates . . . . .	60
4.3.3	Socioeconomic & Demographic Variables . . . . .	61
4.3.4	City Analysis: Park Access & Health . . . . .	61
4.3.5	Tract Level Park Access & Equity . . . . .	63
4.4	Results . . . . .	64
4.4.1	City Analysis: Park Access & Health . . . . .	65
4.4.2	Tract Level Access & Equity . . . . .	66
4.5	Discussion . . . . .	69
4.6	References . . . . .	72
<b>5</b>	<b>Conclusion</b>	<b>76</b>
5.1	Synthesis & Next Steps . . . . .	76
5.2	Broader Impacts . . . . .	78
<b>6</b>	<b>Bibliography</b>	<b>80</b>
<b>7</b>	<b>Appendix</b>	<b>94</b>
7.1	Chapter 3 Appendix . . . . .	94
7.1.1	Twitter API . . . . .	94
7.1.2	Stopwords . . . . .	94
7.1.3	Hashtags . . . . .	94
7.1.4	Happiness Benefit: User Control . . . . .	96
7.1.5	Temporal Analysis by hour of day . . . . .	98
7.2	Chapter 4 Appendix . . . . .	100
7.2.1	500 Cities Data . . . . .	100
7.2.2	500 Cities Validation . . . . .	100
7.2.3	Small-Area Estimation Methodology . . . . .	100



## List of Figures

2.1	San Francisco Recreation and Parks Department facility map . . . . .	13
2.2	Sentiment before, during, and after park visit . . . . .	18
2.3	Change in sentiment between park categories. . . . .	19
2.4	Park tweets vs. baseline tweets wordshift . . . . .	21
2.5	Word frequency patterns before and after park visit. . . . .	22
3.1	Happiness benefit by city . . . . .	42
3.2	Park tweets wordshift . . . . .	44
3.3	ParkScore® and park spending vs. happiness benefit . . . . .	45
3.4	Chicago map and wordshift . . . . .	46
3.5	Happiness benefit by park size, season, and day of the week . . . . .	47
4.1	Scatterplots and regression lines for park access and health outcomes	65
4.2	Regional scatterplots and regression lines for park access and health outcomes . . . . .	66
4.3	Relationships between park access and socioeconomic characteristics of neighborhoods . . . . .	68
7.1	Stop word histogram . . . . .	96
7.2	Happiness benefit by city, user control . . . . .	97
7.3	ParkScore® and park spending vs. happiness benefit, user control . . .	98
7.4	Happiness benefit by hour of day . . . . .	99

## List of Tables

2.1	Park category characteristics . . . . .	17
3.1	Summary of Twitter data for 25 cities . . . . .	38
4.1	Data Sources . . . . .	62
4.2	Data summary by region . . . . .	64
4.3	Regression model results . . . . .	67
7.1	Stop words for 25 cities . . . . .	95

# Chapter 1: Introduction

## 1.1 Problem Statement

The procession of the agricultural, industrial, and digital revolutions has vastly altered the environment and manner in which people spend their time. These massive cultural shifts transformed our society and drew people to cities for economic opportunities. Urbanization has led to over half of the global population living in cities. While cities offer economic and social benefits, urban residents fare worse than rural populations across several health outcomes. Specifically, people who live in cities face declining mental health and increased obesity (McDonald, Beatley, and Elmqvist, 2018). Not only do mental health disorders and obesity cause substantial human suffering, they are responsible for a global economic burden in the billions of dollars (Waters and Graf, 2018).

Spending time in natural environments has shown promise for mitigating these health challenges. In many cultures, the benefits of nature contact have historically been promoted through practices such as *shinrin-yoku*, or forest bathing, in Japan, and *friluftsliv*, or the open-air life, in Scandinavian countries (Buckley and Brough, 2017). Modern lifestyles in cities do not always facilitate time in nature which has led to a reduction in both access and interest among urban residents (Soga and Gaston, 2016). Combined with the increase in digital work, it is not surprising that the average amount of time people spend indoors, sedentary, and looking at screens has increased in recent years (Crnic and Kondo, 2019).

Parks are of particular interest for improving health because they are ubiquitous in modern cities and theoretically accessible by all urban residents. In addition, parks

are typically managed by municipal governments and can be improved or expanded with specific local needs in mind (Larson, Jennings, and Cloutier, 2016). People use parks for social gathering, exercise, and relaxation; these activities are part of the theorized psychological and biological pathways connecting nature contact and health outcomes (Markevyh et al., 2017). With this in mind, park visits have started to be promoted in health campaigns such as park prescription programs (Van den Berg, 2017).

While the scientific consensus on the benefits of nature contact has grown, the promotion of nature contact as a health promoting behavior has not achieved widespread adoption. This may be due to a lack of sufficient parks & recreation department funding, difficulty in translating science to practice, the ambiguity of scientific results & the concept of significance, and the challenges of communicating results to the public (van den Bosch and Nieuwenhuijsen, 2017). Focusing research on illuminating the health benefits of parks is one way to increase the relevance and legibility of research for policymakers, public health experts, and urban dwellers.

## **1.2 Background and Gaps**

Several frameworks have been used to explain the pathways between nature contact and observed changes in health. Typically, these frameworks include a typology of nature contact, predictions for various health changes, and moderating factors that alter the connection between nature contact and observed health impacts (Hartig, Mitchell, de Vries, and Frumkin, 2014; Kuo, 2015). Underlying these pathways are several theories that attempt to explain the biological or psychological mechanisms through which nature contact affects health outcomes. The mechanisms include the activation of specific psychological pathways, enhanced immune responses, greater

physical activity, increased social connection and improved air quality resulting from nature contact (Frumkin et al., 2017). For mental health outcomes, theories include Stress Reduction Theory, which predicts a decrease in physiological stress following nature contact (Ulrich et al., 1991). For physical health outcomes, nature contact may provide increased opportunities for physical activity resulting in reduced obesity or all-cause mortality (Nieuwenhuijsen, Khreis, Triguero-Mas, Gascon, and Dadvand, 2017). Different types or sizes of urban greenspaces may provide different opportunities for these mechanisms to engage; studying details about the green spaces may help uncover which mechanistic pathways are most relevant (Gascon et al., 2015). While reviews have called for direct investigation of these pathways, the methodological approaches typically used to study nature contact vary in their ability to test these mechanisms directly.

A variety of methodological approaches have been used to study the health benefits of nature contact. Here, I briefly describe the strengths and limitations of three main types of methodological approaches: experimental, epidemiological, and experience-based.

Experimental studies are well-suited for measuring physiological and psychological responses to discrete periods of nature contact. Field experiments randomly assign participants to a nature-treatment group and an urban control group. Investigators then measure health or cognitive markers pre- and post-exposure (Bratman, Daily, Levy, and Gross, 2015). These experiments have built a strong evidence base for the positive short-term health and cognitive benefits from nature contact (Krabbendam et al., 2020). Experimental designs usually examine the effects of nature contact following, rather than during exposure, though benefits experienced during exposure may be of interest (Ohly et al., 2016). Translating experimental results to the way people engage with nature in real life can also be challenging.

Epidemiological studies model the relationships between nature contact and health using surveys and geographic data. Compared to experiments, they study the health effects of nature contact over a longer time period with larger sample sizes. By measuring vegetative cover near where people live, these types of studies estimate a continuous metric of nature contact. However, the existence of nearby vegetation does not guarantee its use, and most of these studies have focused on nearby vegetation rather than nearby parks (Boyd, White, Bell, and Burt, 2018; Fong, Hart, and James, 2018). These studies usually have cross-sectional designs which make causal inference challenging; however, they have complemented experimental approaches by pointing to how results may scale to wider populations and a wider range of locations.

Experience-based approaches attempt to capture real-life responses to nature contact by collecting high-resolution individual data (Krabbendam et al., 2020). These studies utilize cellular phones to sample user locations and mood via a mobile app (MacKerron and Mourato, 2013; McEwan et al., 2019). In addition, land cover maps can capture more detailed information on the typology of locations people actually visit. As an alternative to purpose-built mobile applications for research, researchers have used data from public social media posts to estimate exposure to various locations while analyzing the content of those posts (Lim et al., 2018; Roberts et al., 2018). While there may be bias in the types of people willing to use a mobile app or share their lives publicly on social media, experience-based approaches offer a unique way to gain insight into how people engage with nature in the normal course of their day-to-day lives.

All of these approaches have contributed to the growing consensus around the health benefits of nature contact. However, quantifying the health benefits of nature contact, and specifically urban parks, remains a key gap in the literature. Another open question is whether different sizes or types of urban greenspace deliver different

health benefits. Finally, understanding the benefits of park visits beyond a single city is critical - the way people benefit from nature contact may not be uniform across different places.

## 1.3 Research Scope

In this dissertation, I investigate the link between nature contact and health in cities, with a focus on quantifying the benefit of urban parks and understanding how the benefit varies in different geographic contexts. I have organized this dissertation into three independent chapters, each written for publication in a peer-reviewed journal. The chapters work in concert to fill the gaps outlined in the prior section. I have ordered the chapters by increasing geographic scope.

In Chapter 1, I apply an experience-based approach and use Twitter to look at the relationship between nature contact in urban parks and mental well-being in a single city, San Francisco. I analyze the text people write in tweets before, during, and after visits to parks. I then estimate the magnitude and duration of the mental benefits people exhibit in their tweets. I also investigate whether certain types of parks are more beneficial than others.

In Chapter 2, I apply similar methods to the 25 largest cities in the US to test whether Chapter 1's results generalize beyond a single city. I compare estimates of the mental benefits of nature contact across cities and analyze how the benefits of park visits vary with park investment and park size. I also investigate weekly and seasonal variation in the mental benefits of nature contact.

In Chapter 3, I again expand the scope of the analysis, now to the largest 500 cities in the US. In this chapter, I take an epidemiological approach using cross-sectional population data from the US Census, Center for Disease Control, and Trust for Public

Land. I estimate census-tract and city-level park access and look at its relationship with two health outcomes - mental health and obesity. I also analyze census-tract level park access across different socioeconomic and demographic groups to better understand equity of access to urban parks.

I conclude the dissertation by describing how these three chapters improve our understanding of how parks can contribute to improving health outcomes. I summarize how this work elaborates on prior efforts, demonstrates novel methods, and provides an example of how different approaches can be used in concert to enhance our understanding of the impacts of urban parks on human health. After developing some potential avenues for future research, I summarize the broader implications of this research for policymakers, public health officials, and interested citizens.



# Chapter 2: Visitors to urban greenspace have higher sentiment and lower negativity on Twitter

## 2.1 Abstract

With more people living in cities, we are witnessing a decline in exposure to nature. A growing body of research has demonstrated an association between nature contact and improved mood. Here, we used Twitter and the Hedonometer, a word analysis tool, to investigate how sentiment, or the estimated happiness of the words people write, varied before, during, and after visits to San Francisco’s urban park system. We found that sentiment was substantially higher during park visits and remained elevated for several hours following the visit. Leveraging differences in vegetative cover across park types, we explored how different types of outdoor public spaces may contribute to subjective well-being. Tweets during visits to Regional Parks, which are greener and have greater vegetative cover, exhibited larger increases in sentiment than tweets during visits to Civic Plazas and Squares. Finally, we analyzed word frequencies to explore several mechanisms theorized to link nature exposure with mental and cognitive benefits. Negation words such as *no*, *not*, and *don’t* decreased in frequency during visits to urban parks. These results can be used by urban planners and public health officials to better target nature contact recommendations for growing urban populations.

## 2.2 Introduction

There is a growing interest in understanding the connection between mental health and exposure to biodiversity, due to the simultaneous growth of urban areas globally and rising rates of mood disorders (Murray et al., 2012; United Nations, 2014). While cities provide access to significant economic and social opportunities, researchers have identified an urban health penalty that arises from the pace of life, exposure to environmental stressors and chemicals, and disconnect from diverse natural environments in which human evolution occurred (Bettencourt, Lobo, Helbing, Kühnert, and West, 2007; Elmqvist et al., 2015; McDonald, Beatley, and Elmqvist, 2018). Urban greenspace, and specifically urban parks, are a policy instrument that can help reduce the impacts of “nature deficit disorder” (Louv, 2011).

Studies on the mental benefits of nature exposure have typically taken one of two approaches. First, broad studies based on surveys and geographic data have established an association between proximate natural areas and subjective well-being (Hartig, Mitchell, de Vries, and Frumkin, 2014; Maas, Verheij, Groenewegen, de Vries, and Spreeuwenberg, 2006; van den Berg et al., 2016; Wheeler et al., 2015). The Normalized Difference Vegetation Index (NDVI), a proxy for vegetation derived from remotely sensed data, has been used as a measure of neighborhood greenness and is associated with lower levels of depression (Fong, Hart, and James, 2018). High levels of cumulative childhood greenspace exposure were associated with lower risk of developing psychiatric disorders (Engemann et al., 2019). Broad surveys are unable to capture acute exposure events to greenspace and biodiversity, making it challenging to identify the types of natural areas most effective at delivering mental benefits (Bell, Phoenix, Lovell, and Wheeler, 2014). Field experiments address this issue by directly exposing participants to natural areas. For example,

participants walking through natural areas showed improved affect and cognition compared to those walking through urban environments (Bratman, Daily, Levy, and Gross, 2015). In another experiment, individuals who walked in areas with greater biodiversity reported higher levels of subjective well-being (Carrus et al., 2015). Several recent experiments have examined landscape preferences, landscape structure, and biodiversity across a gradient of park types with larger experimental pools (Fischer et al., 2018; Hoyle, Hitchmough, and Jorgensen, 2017; Hoyle, Jorgensen, and Hitchmough, 2019; Martens, Gutscher, and Bauer, 2011; Qiu, Lindberg, and Nielsen, 2013). Here, we combined the strengths of these approaches by following a large group of people making known visits to a range of park types.

Recently, mobile phone applications have been used to conduct ecological momentary assessments, querying users about their mood and environment in real-time (Bakolis et al., 2018; MacKerron and Mourato, 2013; McEwan et al., 2019). In the present study, we use location-enabled data from social media to observe individuals at a level of geographical precision that indicates actual contact with greenspace and biodiversity. Previous studies analyzing tweets in urban greenspace have studied comparative well-being across a city, emotional changes across seasons, and different ways of analyzing the emotional content of tweets (Lim et al., 2018; Plunz et al., 2019; Roberts et al., 2018).

We used the Hedonometer, a word analysis tool that quantifies the sentiment of text based on the happiness values attributed to English words (Dodds and Danforth, 2010; Dodds, Harris, Kloumann, Bliss, and Danforth, 2011). The Hedonometer has been demonstrated to correlate with traditional survey metrics of subjective well-being at the city and state level, including Gallup’s well-being index and United Health Foundation’s America’s health ranking (Mitchell, Frank, Harris, Dodds, and Danforth, 2013). The Hedonometer has also been deployed to analyze the discourse

around climate change following hurricanes (Cody, Reagan, Mitchell, Dodds, and Danforth, 2015). Using the Hedonometer’s sentiment dictionary we asked: (Q1) What is the magnitude and duration of the change in sentiment from visiting urban parks?

Using geo-located tweets allowed us to differentiate between different doses of nature exposure intensity, defined as the quality and quantity of nature itself, as called for in prior work (Shanahan, Fuller, Bush, Lin, and Gaston, n.d.). The San Francisco Recreation and Parks Department classifies their facilities into categories based on park size, design, and amenities. Using official park types along with estimates of park vegetative cover, we investigated how different types of nature exposure are associated with changes in happiness as expressed in tweets. While we don’t measure biodiversity directly, we use NDVI and vegetative cover as proxies for the intensity of nature exposure. We asked: (Q2) What is the association of park type and vegetative cover with the change in sentiment from park visitation?

Complementary theories from psychology and neurobiology suggest several mechanisms connecting nature exposure with mental state (Berto, 2014; Frumkin et al., 2017). The Biophilia hypothesis suggests that humans have an innate affinity for natural environments similar to those in which we evolved (Kellert and Wilson, 1995). More specifically, Stress Reduction Theory (SRT) predicts a decrease in physiological stress following nature contact, resulting in a variety of positive health outcomes (Ulrich et al., 1991). Attention Restoration Theory (ART) predicts that time in nature provides the opportunity to restore directed attention capacity, which results in improved cognition (S. Kaplan, 1995; Ohly et al., 2016). Nature exposure has also been found to correlate with increased prosocial behavior through ‘unselfing’, a shift away from self-interest and towards generosity (Zhang, Piff, Iyer, Koleva, and Keltner, 2014). A recent review of the pathways linking greenspace to health called

for quasi-experimental studies and the assessment of varying exposure types to better explore the mechanisms underlying the mental benefits of nature contact (Markevych et al., 2017). Here, we analyzed word frequency patterns around park visitation to explore the mechanisms driving mental shifts from park visitation. We derived word frequency time series from tweets and asked: (Q3) What do word frequency patterns indicate about the mechanisms driving the change in sentiment from park visitation?

## 2.3 Methods

### 2.3.1 Study Site & Data Collection

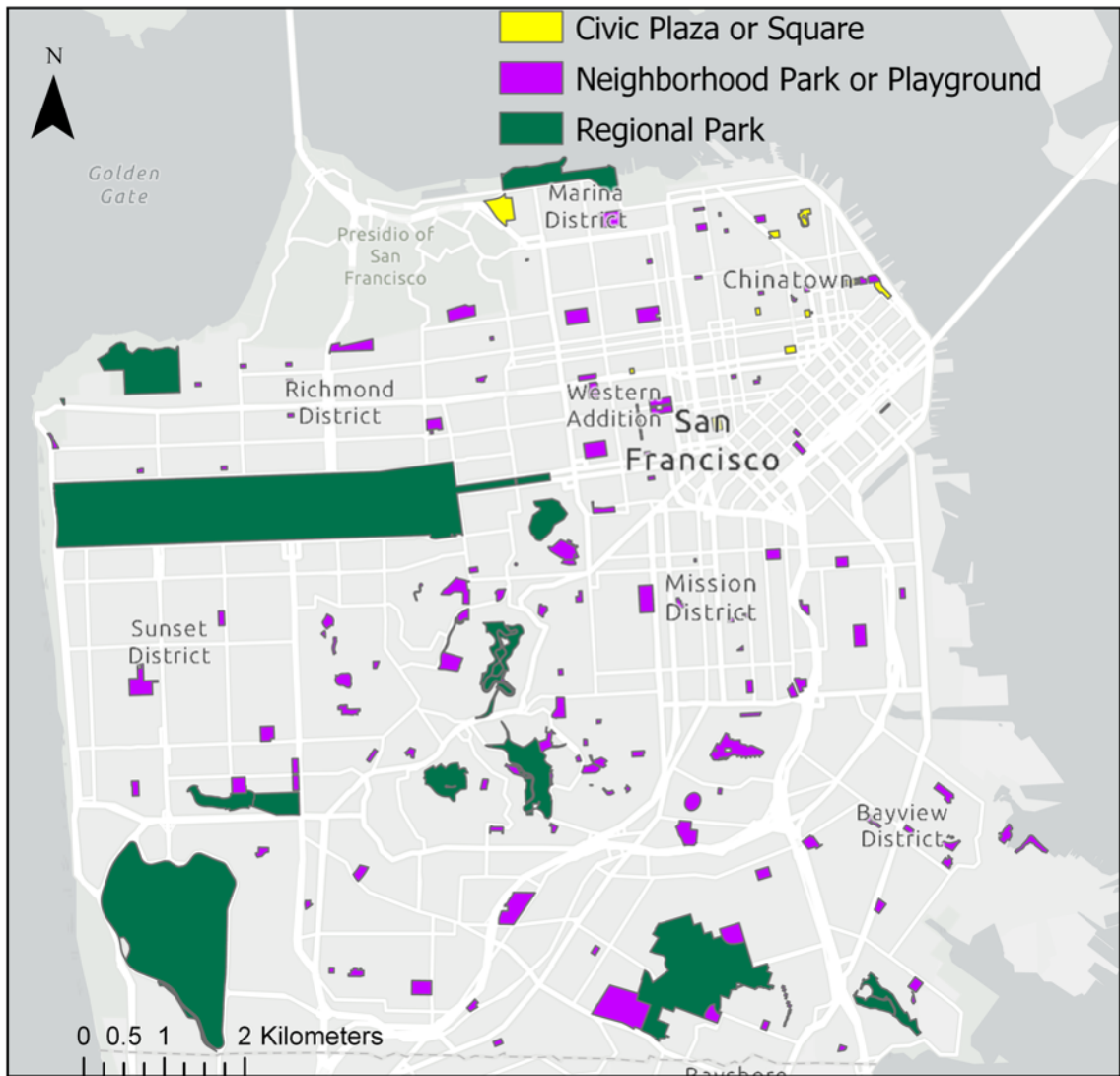
Using Twitter’s streaming Application Programming Interface (“Twitter Streaming API”, 2016), we collected all tweets explicitly geotagged with latitude and longitude originating in the San Francisco, USA (2016 Population Estimate: 871,000) area between May 19, 2016 and August 2, 2016 (roughly 70,000 tweets per day). Although Twitter places a rate limit on the streaming API, our sample came from a relatively small geographic area leading to insignificant errors for tweet collection. We selected San Francisco as a study site due to its diverse park system, which spans more than 220 sites and 3,400 acres. According to the Trust for Public Land, 98.2% of San Francisco’s population live within walking distance of a park and San Francisco has one of the top ranked park systems in the USA (Harnik, Martin, and Barnhart, 2015). Using the `Python` (v.2.7) geographic libraries `Fiona` (v. 1.5.1) and `Shapely` (v. 1.6), we determined which tweets fell within San Francisco Recreation and Parks Department facility boundaries (“Park and Open Space Map”, 2016). Finding tweets inside parks depends on the accuracy of mobile GPS; some of our user pool may have been just outside of parks due to measurement error. San Francisco Recreation

and Parks categorizes their facilities into nine categories, with 94% of Tweets collected located in the following three categories: Regional Parks, Civic Plazas and Squares, and Neighborhood Parks and Playgrounds (Fig. 2.1). These parks were categorized by a professional parks planner according to guidelines determined by San Francisco Recreation and Parks Department.

We constructed a list of Twitter users who had visited at least one park during the study period and used the Twitter API to download their 3,200 most recent tweets. A month later, we updated user histories with any tweets posted since the park visit. We used several heuristics to remove automated bots and businesses from the user sample and additionally removed any individual who made their account private in the period following their park tweet. We also removed users who did not have a park visit tweet in English. Our process resulted in 4,688 user timelines.

### **2.3.2 Tweet Binning**

We saved the following fields for each tweet within a user’s timeline: message identification string, timestamp, text, language, and location. We used tweet timelines as the raw data for all further analysis. We defined a park exposure as the first tweet posted from within a park on a given calendar day. We assigned all other tweets as ‘pre’ or ‘post’ to the closest park exposure tweet before or after, enabling the binning of tweets across users into hourly bins. For example, if a user tweeted in a park a 2PM, and also tweeted at 10:30 AM and 4:15 PM, the user would have tweets in the  $-4$ ,  $0$  (in-park), and  $+3$  bins. If we encountered subsequent tweets that also occurred in parks on a given day, we treated them the same as all other post-park tweets. This avoids the bias of including several consecutive park tweets in the park exposure bin and simplifies the assignment of out of park tweets to the initial exposure. Users had an average of 0.62 in-park tweets within 24 hours of their initial



*Figure 2.1: San Francisco Recreation and Parks Department facility map*

park exposure tweets and 78% of users had no secondary park tweets. By pooling users into relative time bins, we were able to create large enough word samples to apply sentiment analysis.

### 2.3.3 Sentiment Analysis

The Hedonometer includes a sentiment dictionary for 10,022 of the most commonly used English words, merged from four distinct text corpora. The Hedonometer performs favorably compared with other sentiment dictionaries, using a continuum scoring of words with high coverage (Reagan, Danforth, Tivnan, Williams, and Dodds, 2017). Word ratings were calculated by averaging scores from a pool of online crowd-sourced workers at Amazon’s Mechanical Turk (Dodds et al., 2011). The words were rated on a scale from 1 (least happy) to 9 (most happy). For example, *sunshine* has a score of 7.9 and *traffic* has a score of 3.3. Words with scores between 4 and 6 were excluded from the analysis either because they are emotionally neutral (e.g., *at* [4.9], *and* [5.2]) or because they are context dependent (e.g. *church* [5.5], *capitalism* [5.2]). For our study, we also removed any words appearing in the names of San Francisco Parks from the analysis (e.g., *golden* [7.3], *gate* [5.1], and *park* [7.1]). We recognize that words representing natural features such as *beach* [7.9], *tree* [7.1], *grass* [6.5] typically have positive sentiment and hypothesize that the presence of such words indicates awareness of the surrounding environment, which has been connected with a reduction in stress (Ulrich et al., 1991). While the Hedonometer does not take word context or order into account, prior use of the tool has indicated that with a sufficiently large sample size ( $> 1,000$  words), a reliable estimate of text happiness is possible (Reagan et al., 2017). For this reason, we did not measure the happiness of individual tweets or users but instead implemented the pooling procedure described below.



### 2.3.4 Estimating Sentiment

For a set of tweets, we estimated sentiment as the weighted average of word scores using their relative frequencies as weights (Equation 2.1). We generated sentiment time curves by applying the Hedonometer to hourly bins of tweets before, during, and after park exposure. To provide additional statistical support to this approach, we used a bootstrapping procedure. For a given hourly bin, we randomly selected 80% of the tweets without replacement and calculated the pooled sentiment. Performing this procedure 100 times, we derived a range of plausible mean sentiment values for each tweet bin.

$$Sentiment = \frac{\sum_i^n v_i * f_i}{f_i} \quad (2.1)$$

Where  $v_i$  is the sentiment score of the  $i$ th word and  $f_i$  is its frequency in a given text (set of tweets).

To quantify the change in sentiment from exposure to urban greenspace, we compared the sentiment from tweets before park visits and during park exposure. First, we defined a set of baseline tweets. For a given park, these were tweets occurring more than 1 and up to 6 hours prior to tweets posted from that park. We subtracted the baseline sentiment from that of the park exposure tweets. To estimate a plausible interval for change in sentiment, we performed a similar bootstrapping procedure. We selected a random 80% of tweets from both the baseline and park tweets and calculated the difference in their sentiment scores. Performing this operation 100 times, we were able to estimate a mean, variance, and a 95% confidence interval for the mean change in sentiment. Robustness checks were performed to show convergence of this range at 100 runs.

### 2.3.5 Duration Calculation

To estimate the duration of a change in sentiment from visiting a park or set of parks, we defined the baseline set of tweets in the same manner as above. We then performed the following bootstrapping procedure in an iterative manner. We started with the tweets occurring one hour after park exposure and estimated the change in sentiment between the baseline and that hourly bin of tweets. We continued to the next bin if we were able to reject the null hypothesis from the one sample T-Test that the mean of the bootstrapped differences is equal to 0 at the 95% confidence level. The duration of a change in sentiment was the last hourly bin at which we are able to reject this null hypothesis. We performed this analysis with and without in-park tweets that occur after initial park tweets on a given day.

### 2.3.6 Park Classifications and NDVI

To understand how park type relates to the benefits of park exposure, we used the San Francisco Recreation and Parks Department classifications for the 160 parks in which we found tweets during the study period. The vast majority of park acreage and tweet activity occurred in 3 categories: Civic Plaza or Square, Neighborhood Park or Playground, and Regional Park (Fig. 2.1). We grouped tweets posted from each of these park categories to compare the changes in sentiment from baseline across categories.

We also calculated Normalized Difference Vegetation Index (NDVI) for each of the 160 parks in which tweets occurred. We developed an automated method to map vegetation throughout San Francisco using an object-based approach with 1-meter, 4-band National Agricultural Imagery Program (NAIP) data acquired in the summer of 2016 (O’Neil-Dunne, Pelletier, MacFaden, Troy, and Grove, 2009). We segmented

Table 2.1: Park category characteristics

<u>Category</u>	<u>Count</u>	<u>Mean Acres</u>	<u>Mean NDVI</u>	<u>Mean Percent Vegetated</u>
Regional Park	13	609.37	0.21	79.48%
Neighborhood Park or Playground	112	11.54	0.12	63.44%
Civic Plaza or Square	10	8.79	0.06	45.42%

NAIP imagery into image objects using a multiresolution segmentation algorithm (Benz, Hofmann, Willhauck, Lingenfelder, and Heynen, 2004). We computed NDVI for each image object based on the mean near-infrared and red values in the NAIP data (Equation 2.2).

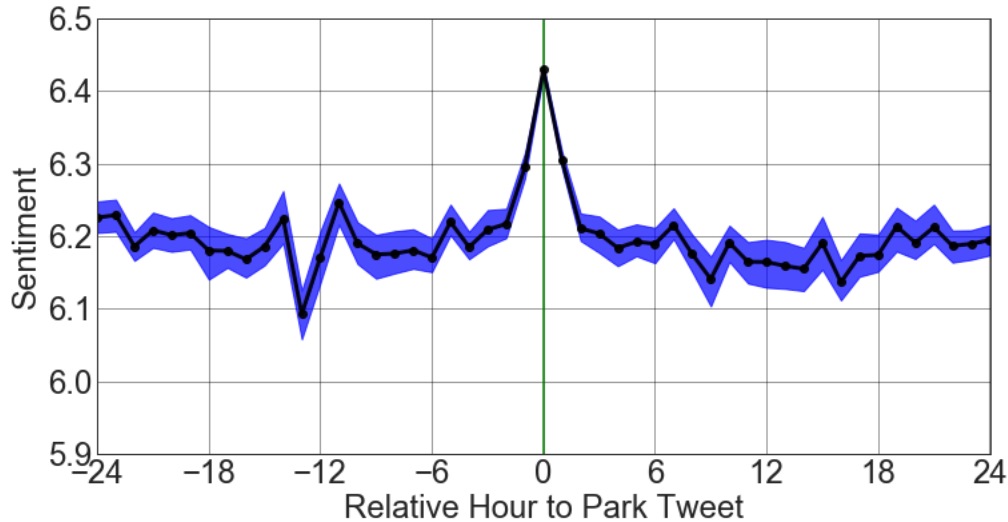
$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (2.2)$$

NIR = near-infrared; RED = visible red.

Using a series of classification and morphology algorithms, we assigned objects to one of two classes: vegetation or non-vegetation. We overlaid these objects, along with their NDVI values, onto the San Francisco park polygons to calculate the percent area with vegetation and mean NDVI for each park, excluding pixels defined as bodies of water from data provided by the SF Department of Public Works (“Water Bodies”, 2019). NDVI scores range from -1 to 1 with higher scores being greener. We report NDVI and Percent Vegetation for the 3 main park categories in Table 2.1. The other park categories (Community Gardens, Concessions, Family Camp, Mini parks, Parkways, and Zoological Gardens) were not large enough to accurately estimate mean NDVI or sentiment based on the number of tweets posted from those spaces.

## 2.4 Results

**Q1: What is the magnitude and duration of the change in sentiment from visiting urban parks?** Tweets posted within parks have a higher sentiment than tweets posted before or after park visits. We depict the sentiment time curve for all users in Figure 2.2, with average sentiment fluctuating between roughly 6.1 and 6.2 outside of park visits. Sentiment reaches 6.43 across all tweets occurring in parks. The immediate hours before and after park exposure also elevated from baseline. The bootstrapped intervals for mean sentiment are narrower around the park exposure because our dataset contains more tweets during those hours than in any individual hour preceding or following the park exposure.

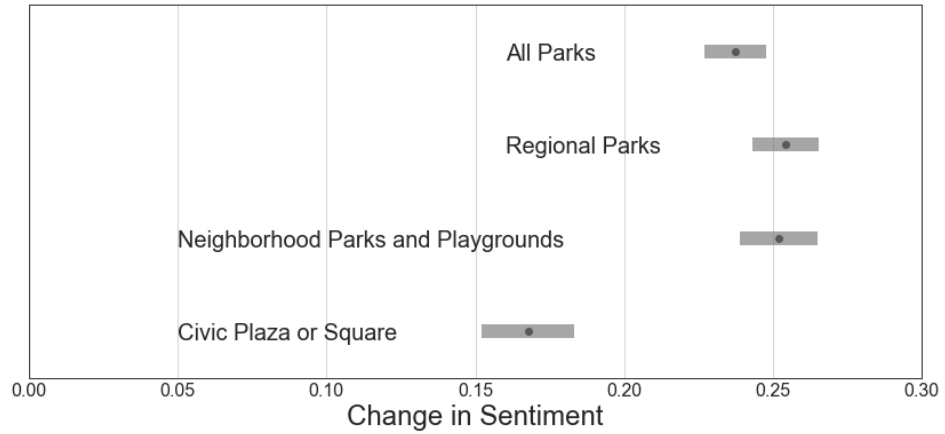


*Figure 2.2: Sentiment before, during, and after park visit. Average sentiment for all user tweets (y-axis), within 24 hours of park exposure, binned by relative hour to in-park tweet (x-axis). The green vertical line represents the tweet in a San Francisco Park, with highest sentiment value. The blue range is full sentiment range from 100 runs of randomly sampling 80% of tweets.*

The mean change in sentiment for all parks is 0.229 (Bootstrapped 95% CI 0.220, 0.238) (Fig. 2.3). As a point of reference, the average day on Twitter in 2016 had a

sentiment of 6.04, and Christmas Day was the happiest day in 2016 with a sentiment of 6.26 (Hedonometer.org, 2020). Thus, across our user pool, tweets during visits to urban parks exhibited a similar increase in sentiment as the jump on Christmas Day for Twitter as a whole.

Across all parks, we estimate the duration of elevated sentiment after initial park tweets. We find that sentiment remains elevated for four hours, compared to a baseline level averaged over the six hours before park visitation. This analysis included tweets inside parks that occurred after the initial exposure to avoid bias of highly active users and to clarify assignment of post park tweets to an initial park exposure (see methods). To calculate an even more conservative estimate of elevated sentiment, we repeated this analysis without those tweets, resulting in an estimated duration of one hour. We thus expect the duration to fall somewhere between one and four hours.



*Figure 2.3: Change in sentiment between park categories. The horizontal axis is estimated change in sentiment between baseline and in-park tweets. Ranges are 95% Confidence Intervals from 100 runs of bootstrap process. Dots are mean change in sentiment. Park categories are as defined by San Francisco Recreation and Parks Department.*

**Q2: What is the association of park type and vegetative cover with the change in sentiment from park visitation?** Regional Parks exhibit the highest mean change in sentiment of 0.264 (0.251, 0.276). Neighborhood Parks and

Playgrounds have a moderate mean change in sentiment of 0.219 (0.199,0.240). Civic Plaza or Squares have the lowest change in sentiment of 0.163 (0.143, 0.181) (Fig 2.3). Confidence interval limits do not overlap for any pair of park types, indicating significant differences among them. The differences in mean sentiment among park types correspond to differences in size, NDVI and vegetative cover (Table 2.1).

**Q3: What do word frequency patterns indicate about the mechanisms driving the change in sentiment from park visitation?** Tweets in parks have higher sentiment than tweets prior to park visitation due to positive words with higher frequency, such as *beach*, *beautiful*, *festival*, *happy*, *young*, *fun*, and negative words with lower frequency, such as *not*, *no*, *don't*, *can't*, and *wait* (Fig. 2.4). Of specific interest, negation words such as *not* and *don't* fluctuate before and after park exposure but exhibit a marked drop (45% and 47%) in and around the park exposure (Fig. 2.5). The word *beautiful* exhibits the opposite pattern, fluctuating around a baseline and then roughly doubling in frequency during park exposure (Fig. 2.5C). Finally, we examine the first-person word *me* which has a neutral sentiment (and is not included in the sentiment scores above). Use of *me* drops 38% from its mean use level during park visits (Fig. 2.5D).

## 2.5 Discussion

In our study, tweets posted from urban nature were happier by roughly 0.23 points on the Hedonometer scale from baseline. This increase in sentiment is equivalent to that of Christmas Day for Twitter as a whole in the same year (Hedonometer.org, 2020). Our analysis of duration suggests that elevated sentiment lasts for between 1 and 4 hours following an initial park tweet. The recent Urban Mind study found a similar duration for their week-long study on roughly 100 users self-reporting their

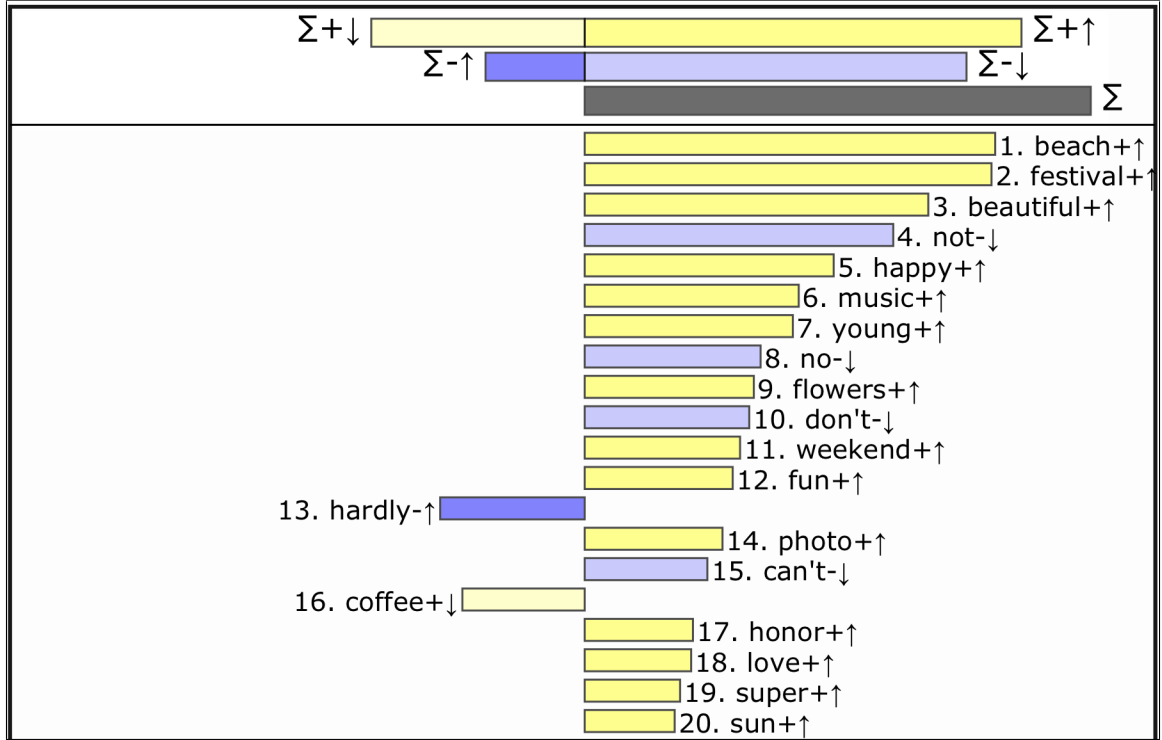


Figure 2.4: Park tweets vs. baseline tweets wordshift. This figure shows the words driving the difference between park and baseline tweets, in order of decreasing contribution to the difference in sentiment. The right side represents the park tweets, with a mean sentiment of 6.43. The left side represents tweets 1-6 hours preceding the park tweets, with a mean sentiment of 6.20. Purple bars represent words  $\leq 4$  (with - symbol) on the Hedonometer scale. Yellow bars represent words  $\geq 6$  (with + symbol) on the Hedonometer scale. Arrows indicate whether a word was more or less frequent within that set of tweets compared to the other text. For example, beach is a positive word (light purple) with higher frequency in park tweets that contributes to increased sentiment of that set. Not is a negative word (light purple) that appears less frequently in that set, resulting in a higher relative sentiment score compared to baseline.

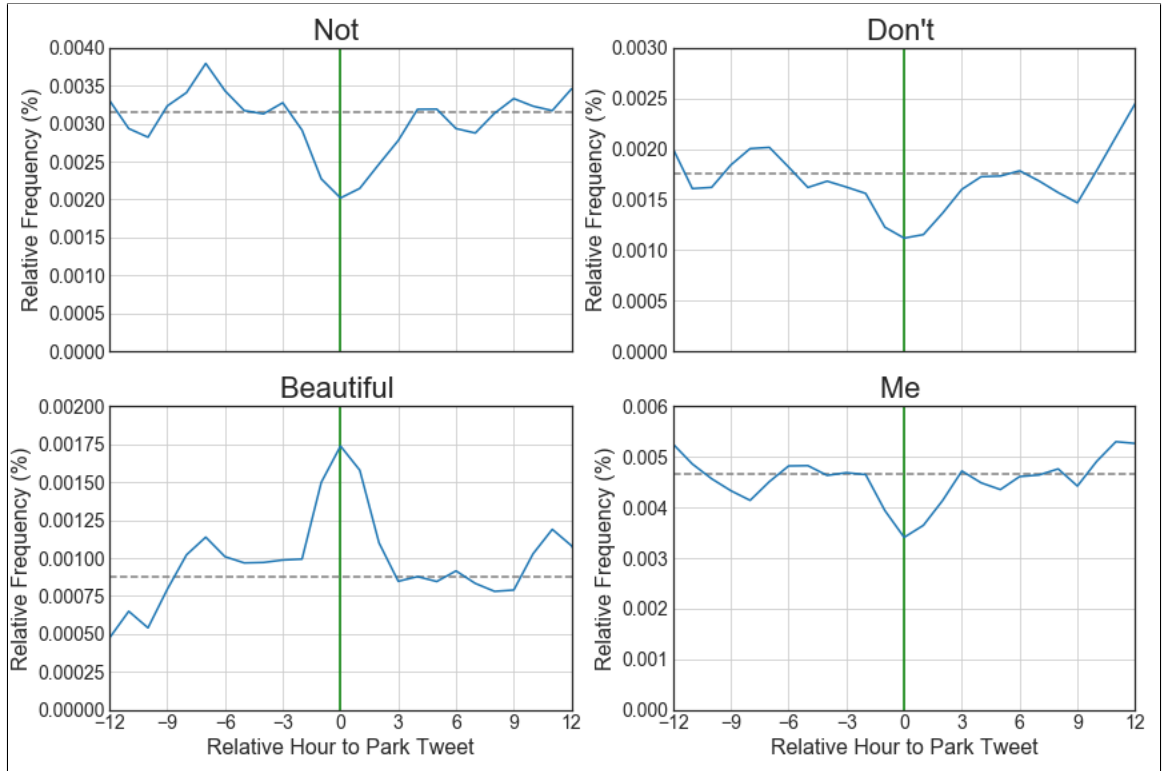


Figure 2.5: Word frequency patterns before and after park visit. X-axis depicts hourly tweet bins from 12 hours before to 12 hours after in-park tweet, which is represented by green line. Y-axis ranges are scaled for each word's relative frequency. Relative frequencies (blue-lines) are smoothed as moving averages over 3 hours. Grey dashed line is mean frequency for entire 24-hour period around park visit.



happiness in different environments (Bakolis et al., 2018). Interestingly, sentiment begins to increase from baseline in the hour preceding the in-park tweets (Fig. 2.2). Possible explanations for this trend are anticipation for the park visit, meeting friends on the way, or perhaps relief due to leaving work and heading to a more enjoyable location. Recent work found that the emotion of anticipation increases in greenspaces (Lim et al., 2018); further investigation is merited to better understand the temporal dynamics of anticipation and its relationship with nature contact.

Tweets located in Regional Parks exhibited the strongest increase in sentiment followed by tweets in Neighborhood Parks and Playgrounds and then Civic Plazas and Squares. There are several plausible explanations for the greater sentiment increase occurring in Regional and Neighborhood Parks. Regional Parks have greater vegetative cover and may offer more opportunities for nature contact and exposure to biodiversity compared to Civic Plazas and Squares (Table 1). The greater vegetative and floral diversity found in the larger Regional Parks may be playing a role as indicated by flowers appearing in Figure 2.4, supported by prior research on the most salient features of landscapes (Hoyle et al., 2017). Alternatively, the large size of Regional Parks may be providing greater restorative capacities through a more distinct separation from the urban environment. Neighborhood Parks and Civic Plazas are close in size but also exhibit a significant difference in sentiment increase, suggesting that park size is not the only factor at play. The three park classes offer different amenities and activity types, which may also be contributing to the differences in sentiment. A recent review summarized the range of scales at which biodiversity can be measured inside of parks - from vegetated versus non-vegetated to genetic diversity - and suggested several directions for future research on the people-biodiversity interface (Botzat, Fischer, and Kowarik, 2016).

The roles of exercise and socialization can be difficult to separate from the direct

contributions of nature to enhanced subjective well-being (Ambrey, 2016). Regional Parks may be more amenable to physical activity, although our analysis of words does not suggest that physical activity related terms are driving the elevated happiness of the in-park tweets (Fig. 2.4). Technologies such as heart-rate monitors may provide new opportunities for distinguishing the benefits of outdoor exercise from the benefits of nature exposure per se. Differences in social interactions across park types may also be contributing to variation in sentiment. Civic Plazas, which tend to be paved and more centrally located, represent an outdoor, public gathering space where people go to socialize in their time away from work. Our results indicate that Regional and Neighborhood Parks are more restorative spaces than Civic Plazas, and that nature per se is potentially playing a role in delivering mental benefits to park visitors. While we are unable to measure how much time is spent in a park following a tweet, it is plausible that visits to Regional Parks are longer than visits to the other park classes. Alternative approaches such as Ecological Momentary Assessments or time use surveys may be more effective at capturing the duration of park visits. Recent work has suggested that at least 120 minutes of weekly nature exposure lead to enhanced self-reported health and well-being (White et al., 2019). Future analyses should also look to directly compare nature contact with indoor activities (e.g., museum visits), but these comparisons are beyond the scope of this paper. Several reviews have summarized the growing body of work on nature contact and set a research agenda for building a more nuanced understanding of the relationship between nature contact and health (Frumkin et al., 2017; Hartig et al., 2014).

The mechanisms through which urban nature exposure improves mental health are still being investigated. Green Mind Theory, a recent synthesis of proposed pathways, suggests that the negativity bias of the brain - which may have been evolutionarily advantageous - is constantly activated by the stressors of modern life

(Pretty, Rogerson, and Barton, 2017). In our analysis, park visitation coincides with a decrease in words such as *no*, *don't*, and *never* (Fig. 2.4). These words, known as negations, are associated with focused, analytical thinking (Pennebaker, 2011). The decrease in negation frequency may provide support for Attention Restoration Theory, which links nature exposure with the experience of soft fascination and can result in improved cognition (R. Kaplan and Kaplan, 1989; Ohly et al., 2016). Alternatively, the increase in frequency of words such as *beautiful*, *fun*, and *enjoy* during park exposure suggest that individuals may be experiencing an increase in positive emotions and a reduction in stress, as predicted by Stress Reduction Theory (Berto, 2014). While the words *I* and *me* do not have an impact on our quantitative analysis due to their neutral sentiment values, there is a distinct decrease in use of these first-person pronouns during park exposure (Fig. 2.5D). This pattern supports prior work describing nature exposure as an opportunity to shift from an individual to collective mental frame, potentially leading to pro-social behavior (Zhang et al., 2014).

There are also limitations of using Twitter as a platform and we acknowledge our sample of users willing to geolocate may differ from the general population. In 2016, 24% of online adults were active Twitter users, albeit with a slight overrepresentation by younger Americans (Greenwood, Perrin, and Duggan, 2016). Due to the difficulty of extracting this information from Twitter profiles, we were unable to look at how age, gender, and education levels interact with changes in sentiment from park visitation. There is significant variation in how individuals experience and relate to nature; future work should attempt to understand how individual traits mediate the effects of visiting urban greenspace (Gascon et al., 2015). We also recognize that different socioeconomic groups and culturally diverse populations respond differently to conceptions of nature and call for further work teasing apart how varied groups

respond to nature contact (Fischer et al., 2018; Frumkin et al., 2017; Maas et al., 2006). Furthermore, cultures in distinct climates will likely demonstrate different relationships with nature exposure - responses to nature contact will likely manifest very differently in a tropical climate compared to San Francisco (Saw, Lim, and Carrasco, 2015).

In this study, we quantified the change in expressed sentiment associated with visits to urban nature by thousands of individuals. In our sample, individuals tweet happier words while visiting parks, and continue to use happier words for several hours following their visit. Tweets posted in Regional Parks, which are larger and greener, are happier than tweets posted in the smaller and less vegetated Civic Plazas and Squares. Based on our word frequency analysis, improved Twitter sentiment from park visits is driven in part by a decline in negative thinking. Our study deepens the evidence base for the mental benefits provided by nature contact in urban areas. As we continue to uncover the psychological mechanisms from nature contact, we can better inform public health policy and target park planning and design to maximize these benefits.

Urban parks can provide restorative environments for people as well as refuge for biodiversity. The benefits of urban nature include many ecosystem services beyond the scope of this study such as storm-water retention and air purification (Elmqvist et al., 2015). Conserving natural spaces and protecting mental health are not typically discussed in the same policy arenas; however, research further linking health with urban greenspace and biodiversity protection can help planners and public health officials build new strategies that support both goals. We suggest building or expanding parks near populations with limited access to greenspace and targeting funds toward the most effective types of parks for mental benefits. With most of the planet’s population now living in cities, we must find ways to bring nature to them

in a way that supports both biodiversity and human health.

## 2.6 References

- Ambrey, C. L. (2016). An investigation into the synergistic wellbeing benefits of greenspace and physical activity: Moving beyond the mean. *Urban Forestry & Urban Greening*, 19, 7–12.  
DOI: [10.1016/j.ufug.2016.06.020](https://doi.org/10.1016/j.ufug.2016.06.020)
- Bakolis, I., Hammoud, R., Smythe, M., Gibbons, J., Davidson, N., Tognin, S., & Mechelli, A. (2018). Urban mind: Using smartphone technologies to investigate the impact of nature on mental well-being in real time. *BioScience*, 68(2), 134–145.  
DOI: [10.1093/biosci/bix149](https://doi.org/10.1093/biosci/bix149)
- Bell, S. L., Phoenix, C., Lovell, R., & Wheeler, B. W. (2014). Green space, health and wellbeing: making space for individual agency. *Health & Place*, 30, 287–292.  
DOI: [10.1016/j.healthplace.2014.10.005](https://doi.org/10.1016/j.healthplace.2014.10.005)
- Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., & Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3-4), 239–258.  
DOI: [10.1016/j.isprsjprs.2003.10.002](https://doi.org/10.1016/j.isprsjprs.2003.10.002)
- Berto, R. (2014). The Role of Nature in Coping with Psycho-Physiological Stress: A Literature Review on Restorativeness. *Behavioral Sciences*, 4(4), 394–409.  
DOI: [10.3390/bs4040394](https://doi.org/10.3390/bs4040394)
- Bettencourt, L. M., Lobo, J., Helbing, D., Kühnert, C., & West, G. B. (2007). Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the national academy of sciences*, 104(17), 7301–7306.  
DOI: [10.1073/pnas.0610172104](https://doi.org/10.1073/pnas.0610172104)
- Botzat, A., Fischer, L. K., & Kowarik, I. (2016). Unexploited opportunities in understanding liveable and biodiverse cities. A review on urban biodiversity perception and valuation. *Global Environmental Change*, 39, 220–233.  
DOI: [10.1016/j.gloenvcha.2016.04.008](https://doi.org/10.1016/j.gloenvcha.2016.04.008)
- Bratman, G., Daily, G. C., Levy, B. J., & Gross, J. J. (2015). The benefits of nature experience: Improved affect and cognition. *Landscape and Urban Planning*, 138, 41–50.  
DOI: [10.1016/j.landurbplan.2015.02.005](https://doi.org/10.1016/j.landurbplan.2015.02.005)

- Carrus, G., Scopelliti, M., Laforteza, R., Colangelo, G., Ferrini, F., Salbitano, F., ... Sanesi, G. (2015). Go greener, feel better? The positive effects of biodiversity on the well-being of individuals visiting urban and peri-urban green areas. *Landscape and Urban Planning*, 134, 221–228.  
DOI: [10.1016/j.landurbplan.2014.10.022](https://doi.org/10.1016/j.landurbplan.2014.10.022)
- Cody, E. M., Reagan, A. J., Mitchell, L., Dodds, P. S., & Danforth, C. M. (2015). Climate change sentiment on Twitter: An unsolicited public opinion poll. *PLOS ONE*, 10(8), e0136092.  
DOI: [10.1371/journal.pone.0136092](https://doi.org/10.1371/journal.pone.0136092)
- Dodds, P. S., & Danforth, C. M. (2010). Measuring the happiness of large-scale written expression: Songs, blogs, and presidents. *Journal of Happiness Studies*, 11(4), 441–456.  
DOI: [10.1007/s10902-009-9150-9](https://doi.org/10.1007/s10902-009-9150-9)
- Dodds, P. S., Harris, K. D., Kloumann, I. M., Bliss, C. A., & Danforth, C. M. (2011). Temporal patterns of happiness and information in a global social network: Hedonometrics and twitter. *PloS one*, 6(12).  
DOI: [10.1371/journal.pone.0026752](https://doi.org/10.1371/journal.pone.0026752)
- Elmqvist, T., Setälä, H., Handel, S. N., van der Ploeg, S., Aronson, J., Blignaut, J. N., ... de Groot, R. (2015). Benefits of restoring ecosystem services in urban areas. *Current Opinion in Environmental Sustainability*, 14, 101–108.  
DOI: [10.1016/j.cosust.2015.05.001](https://doi.org/10.1016/j.cosust.2015.05.001)
- Engemann, K., Pedersen, C. B., Arge, L., Tsirogiannis, C., Mortensen, P. B., & Svenning, J.-C. (2019). Residential green space in childhood is associated with lower risk of psychiatric disorders from adolescence into adulthood. *Proceedings of the national academy of sciences*, 116(11), 5188–5193.  
DOI: [10.1073/pnas.1807504116](https://doi.org/10.1073/pnas.1807504116)
- Fischer, L. K., Honold, J., Cvejić, R., Delshammar, T., Hilbert, S., Laforteza, R., ... Kowarik, I. (2018). Beyond green: Broad support for biodiversity in multicultural European cities. *Global Environmental Change*, 49(February), 35–45.  
DOI: [10.1016/j.gloenvcha.2018.02.001](https://doi.org/10.1016/j.gloenvcha.2018.02.001)
- Fong, K. C., Hart, J. E., & James, P. (2018). A review of epidemiologic studies on greenness and health: updated literature through 2017. *Current Environmental Health Reports*, 1–11.  
DOI: [10.1007/s40572-018-0179-y](https://doi.org/10.1007/s40572-018-0179-y)
- Frumkin, H., Bratman, G. N., Breslow, S. J., Cochran, B., Kahn Jr, P. H., Lawler, J. J., ... Wood, S. A. (2017). Nature contact and human health: A research

- agenda. *Environmental Health Perspectives*, 125(7), 1–18.  
DOI: [10.1289/EHP1663](https://doi.org/10.1289/EHP1663)
- Gascon, M., Triguero-Mas, M., Martínez, D., Dadvand, P., Forns, J., Plasència, A., & Nieuwenhuijsen, M. (2015). Mental health benefits of long-term exposure to residential green and blue spaces: a systematic review. *International Journal of Environmental Research and Public Health*, 12(4), 4354–4379.  
DOI: [10.3390/ijerph120404354](https://doi.org/10.3390/ijerph120404354)
- Greenwood, S., Perrin, A., & Duggan, M. (2016). Demographics of social media users in 2016. <https://www.pewresearch.org/>.
- Harnik, P., Martin, A., & Barnhart, K. (2015). 2015 City Park Facts. <https://www.tpl.org>. Trust for Public land.
- Hartig, T., Mitchell, R., de Vries, S., & Frumkin, H. (2014). Nature and Health. *Annual Review of Public Health*, 35(1), 207–228.  
DOI: [10.1146/annurev-publhealth-032013-182443](https://doi.org/10.1146/annurev-publhealth-032013-182443)
- Hedonometer.org. (2020). Hedonometer, from The Vermont Complex Systems Center & Computational Story Lab. <http://hedonometer.org/>.
- Hoyle, H., Hitchmough, J., & Jorgensen, A. (2017). All about the ‘wow factor’? The relationships between aesthetics, restorative effect and perceived biodiversity in designed urban planting. *Landscape and Urban Planning*, 164(2017), 109–123.  
DOI: [10.1016/j.landurbplan.2017.03.011](https://doi.org/10.1016/j.landurbplan.2017.03.011)
- Hoyle, H., Jorgensen, A., & Hitchmough, J. D. (2019). What determines how we see nature? Perceptions of naturalness in designed urban green spaces. *People and Nature*, (April).  
DOI: [10.1002/pan3.19](https://doi.org/10.1002/pan3.19)
- Kaplan, R., & Kaplan, S. (1989). *The experience of nature: A psychological perspective*. CUP Archive.
- Kaplan, S. (1995). The restorative benefits of nature: Toward an integrative framework. *Journal of Environmental Psychology*, 15(3), 169–182.  
DOI: [10.1016/0272-4944\(95\)90001-2](https://doi.org/10.1016/0272-4944(95)90001-2)
- Kellert, S. R., & Wilson, E. O. (1995). *The biophilia hypothesis*. Island Press.
- Lim, K. H., Lee, K. E., Kendal, D., Rashidi, L., Naghizade, E., Winter, S., & Vasardani, M. (2018). The grass is greener on the other side: Understanding the effects of green spaces on Twitter user sentiments. In *Companion of the the web conference 2018 on the web conference 2018 - www '18* (pp. 275–282).  
DOI: [10.1145/3184558.3186337](https://doi.org/10.1145/3184558.3186337)

- Louv, R. (2011). *The nature principle: Human restoration and the end of nature-deficit disorder*. Algonquin Books.
- Maas, J., Verheij, R. A., Groenewegen, P. P., de Vries, S., & Spreeuwenberg, P. (2006). Green space, urbanity, and health: how strong is the relation? *Journal of Epidemiology and Community Health*, 60(7), 587–592.  
DOI: [10.1136/jech.2005.043125](https://doi.org/10.1136/jech.2005.043125)
- MacKerron, G., & Mourato, S. (2013). Happiness is greater in natural environments. *Global Environmental Change*, 23(5), 992–1000.  
DOI: [10.1016/j.gloenvcha.2013.03.010](https://doi.org/10.1016/j.gloenvcha.2013.03.010)
- Markevych, I., Schoierer, J., Hartig, T., Chudnovsky, A., Hystad, P., Dzhambov, A. M., ... Fuertes, E. (2017). Exploring pathways linking greenspace to health: Theoretical and methodological guidance. *Environmental Research*, 158(June), 301–317.  
DOI: [10.1016/j.envres.2017.06.028](https://doi.org/10.1016/j.envres.2017.06.028)
- Martens, D., Gutscher, H., & Bauer, N. (2011). Walking in “wild” and “tended” urban forests: The impact on psychological well-being. *Journal of Environmental Psychology*, 31(1), 36–44.  
DOI: [10.1016/j.jenvp.2010.11.001](https://doi.org/10.1016/j.jenvp.2010.11.001)
- McDonald, R. I., Beatley, T., & Elmqvist, T. (2018). The green soul of the concrete jungle: the urban century, the urban psychological penalty, and the role of nature. *Sustainable Earth*, 1x'1(1), 3.  
DOI: [10.1186/s42055-018-0002-5](https://doi.org/10.1186/s42055-018-0002-5)
- McEwan, K., Richardson, M., Brindley, P., Sheffield, D., Tait, C., Johnson, S., ... Ferguson, F. J. (2019). Shmapped: development of an app to record and promote the well-being benefits of noticing urban nature. *Translational Behavioral Medicine*, 1–11.  
DOI: [10.1093/tbm/ibz027](https://doi.org/10.1093/tbm/ibz027)
- Mitchell, L., Frank, M. R., Harris, K. D., Dodds, P. S., & Danforth, C. M. (2013). The geography of happiness: Connecting twitter sentiment and expression, demographics, and objective characteristics of place. *PLoS ONE*, 8(5), e64417.  
DOI: [10.1371/journal.pone.0064417](https://doi.org/10.1371/journal.pone.0064417)
- Murray, C. J., Vos, T., Lozano, R., Naghavi, M., Flaxman, A. D., Michaud, C., ... Abdalla, S., et al. (2012). Disability-adjusted life years (dalys) for 291 diseases and injuries in 21 regions, 1990–2010: A systematic analysis for the global burden of disease study 2010. *The lancet*, 380(9859), 2197–2223.  
DOI: [10.1016/S0140-6736\(12\)61689-4](https://doi.org/10.1016/S0140-6736(12)61689-4)
- O'Neil-Dunne, J., Pelletier, K., MacFaden, S., Troy, A., & Grove, J. (2009). Object-based high-resolution land-cover mapping. In *2009 17th international conference*



- on geoinformatics (pp. 1–6).  
DOI: [10.1109/GEOINFORMATICS.2009.5293435](https://doi.org/10.1109/GEOINFORMATICS.2009.5293435)
- Ohly, H., White, M. P., Wheeler, B. W., Bethel, A., Ukoumunne, O. C., Nikolaou, V., & Garside, R. (2016). Attention Restoration Theory: A systematic review of the attention restoration potential of exposure to natural environments. *Journal of Toxicology and Environmental Health, Part B*, 19(7), 305–343.  
DOI: [10.1080/10937404.2016.1196155](https://doi.org/10.1080/10937404.2016.1196155)
- Park and Open Space Map. (2016). <https://data.sfgov.org/>. San Francisco Recreation & Parks.
- Pennebaker, J. W. (2011). *The Secret Life of Pronouns: What our Words Say about us*.  
DOI: [10.1093/llc/fqt006](https://doi.org/10.1093/llc/fqt006)
- Plunz, R. A., Zhou, Y., Carrasco Vintimilla, M. I., Mckeown, K., Yu, T., Uguccioni, L., & Sutto, M. P. (2019). Twitter sentiment in New York City parks as measure of well-being. *Landscape and Urban Planning*, 189(May), 235–246.  
DOI: [10.1016/j.landurbplan.2019.04.024](https://doi.org/10.1016/j.landurbplan.2019.04.024)
- Pretty, J., Rogerson, M., & Barton, J. (2017). Green mind theory: How brain-body-behaviour links into natural and social environments for healthy habits. *International Journal of Environmental Research and Public Health*, 14(7), 706.  
DOI: [10.3390/ijerph14070706](https://doi.org/10.3390/ijerph14070706)
- Qiu, L., Lindberg, S., & Nielsen, A. B. (2013). Is biodiversity attractive?—On-site perception of recreational and biodiversity values in urban green space. *Landscape and Urban Planning*, 119, 136–146.  
DOI: [10.1016/j.landurbplan.2013.07.007](https://doi.org/10.1016/j.landurbplan.2013.07.007)
- Reagan, A. J., Danforth, C. M., Tivnan, B., Williams, J. R., & Dodds, P. S. (2017). Sentiment analysis methods for understanding large-scale texts: a case for using continuum-scored words and word shift graphs. *EPJ Data Science*, 6(1), 28.  
DOI: [10.1140/epjds/s13688-017-0121-9](https://doi.org/10.1140/epjds/s13688-017-0121-9)
- Roberts, H., Resch, B., Sadler, J., Chapman, L., Petutschnig, A., & Zimmer, S. (2018). Investigating the emotional responses of individuals to urban green space using twitter data: A critical comparison of three different methods of sentiment analysis. *Urban Planning*, 3(1), 21.  
DOI: [10.17645/up.v3i1.1231](https://doi.org/10.17645/up.v3i1.1231)
- Saw, L. E., Lim, F. K. S., & Carrasco, L. R. (2015). The relationship between natural park usage and happiness does not hold in a tropical city-state. *PLOS ONE*, 10(7), e0133781.  
DOI: [10.1371/journal.pone.0133781](https://doi.org/10.1371/journal.pone.0133781)

- Shanahan, D. F., Fuller, R. A., Bush, R., Lin, B. B., & Gaston, K. J. (n.d.). The health benefits of urban nature: how much do we need? *BioScience*, (5), 476–485.  
DOI: [10.1093/biosci/biv032](https://doi.org/10.1093/biosci/biv032)
- Twitter Streaming API. (2016). <https://developer.twitter.com>.
- Ulrich, R. S., Simons, R. F., Losito, B. D., Fiorito, E., Miles, M. A., & Zelson, M. (1991). Stress recovery during exposure to natural and urban environments. *Journal of Environmental Psychology*, 11(3), 201–230.  
DOI: [10.1016/S0272-4944\(05\)80184-7](https://doi.org/10.1016/S0272-4944(05)80184-7)
- United Nations. (2014). World Urbanization Prospects: The 2014 Revision-Highlights. <https://www.un.org/en/development>. United Nations.
- van den Berg, M., van Poppel, M., van Kamp, I., Andrusaityte, S., Balseviciene, B., Cirach, M., ... Maas, J. (2016). Visiting green space is associated with mental health and vitality: A cross-sectional study in four european cities. *Health & Place*, 38, 8–15.  
DOI: [10.1016/j.healthplace.2016.01.003](https://doi.org/10.1016/j.healthplace.2016.01.003)
- Water Bodies. (2019). <https://data.sfgov.org/>. San Francisco Department of Public Works, San Francisco, CA.
- Wheeler, B. W., Lovell, R., Higgins, S. L., White, M. P., Alcock, I., Osborne, N. J., ... Depledge, M. H. (2015). Beyond greenspace: an ecological study of population general health and indicators of natural environment type and quality. *International Journal of Health Geographics*, 14(1), 17.  
DOI: [10.1186/s12942-015-0009-5](https://doi.org/10.1186/s12942-015-0009-5)
- White, M. P., Alcock, I., Grellier, J., Wheeler, B. W., Hartig, T., Warber, S. L., ... Fleming, L. E. (2019). Spending at least 120 minutes a week in nature is associated with good health and wellbeing. *Scientific Reports*, 9(1), 7730.  
DOI: [10.1038/s41598-019-44097-3](https://doi.org/10.1038/s41598-019-44097-3)
- Zhang, J. W., Piff, P. K., Iyer, R., Koleva, S., & Keltner, D. (2014). An occasion for unselfing: Beautiful nature leads to prosociality. *Journal of Environmental Psychology*, 37, 61–72.  
DOI: [10.1016/j.jenvp.2013.11.008](https://doi.org/10.1016/j.jenvp.2013.11.008)

# Chapter 3: Gauging the happiness benefit of US urban parks through Twitter

## 3.1 Abstract

The relationship between nature contact and mental well-being has received increasing attention in recent years. While a body of evidence has accumulated demonstrating a positive relationship between time in nature and mental well-being, there have been few studies comparing this relationship in different locations over long periods of time. In this study, we estimate a happiness benefit, the difference in expressed happiness between in- and out-of-park tweets, for the 25 largest cities in the US by population. People write happier words during park visits when compared with non-park user tweets collected around the same time. While the words people write are happier in parks on average and in most cities, we find considerable variation across cities. Tweets are happier in parks at all times of the day, week, and year, not just during the weekend or summer vacation. Across all cities, we find that the happiness benefit is highest in parks larger than 100 acres. Overall, our study suggests the happiness benefit associated with park visitation is on par with US holidays such as Thanksgiving and New Year’s Day.

## 3.2 Introduction

Human health and well-being depends on the environment in which we live. Most people now live in cities, places defined by built infrastructure where remnant nature and vegetation is planned or managed. Urban greenspace, and specifically urban parks, can provide opportunities to reduce the impacts of the “urban health penalty,”

which includes higher levels of stress and depression in urban residents (McDonald, Beatley, and Elmqvist, 2018). Nature contact is theorized to promote mental health through several complementary pathways including the physiological reduction of stress and the opportunity to restore mental fatigue (Berto, 2014). These pathways have been explored using a dose-response framework which describe the duration, frequency, and intensity of nature contact. Researchers have used experimental, epidemiological, and experience-based approaches to build a consensus around the mental health benefits of urban nature (Krabbendam et al., 2020; Van den Berg et al., 2015). However, there are several questions remaining about how the benefits of nature contact vary across cities (Frumkin et al., 2017).

Nature contact occurs within a specific context, and the ability of urban residents to benefit from greenspace may vary geographically. For example, four large cities showed different effect sizes for their associations between nearby nature and well-being (Taylor, Hahs, and Hochuli, 2018). Larger studies have proved difficult however; a recent review of studies on mental well-being and greenspace in adults was unable to conduct a meta-analysis across locations due to methodological heterogeneity (Houlden, Weich, de Albuquerque, Jarvis, and Rees, 2018). Methods such as Ecological Momentary Assessments and data from social media offer the opportunity to study nature contact at wider spatial scales. Prior work using data from Twitter has established that on average, in-park tweets are happier than tweets originating elsewhere in cities (Schwartz, Dodds, O’Neil-Dunne, Danforth, and Ricketts, 2019). However, it has not been shown that this pattern will hold across a wider selection of cities. The ability to access and enjoy nature is heterogeneous across cities — urban park systems vary widely in quality and investment (Rigolon, Browning, and Jennings, 2018). A recent study found that county area park expenditures were associated with better self-rated health (Mueller, Park, and Mowen, 2019). We hypothesize that cities

with higher levels of investment in parks will provide greater benefits to the mental well-being of park visitors. Understanding inter-city variation in the mental health benefits of nature contact can inform urban planning and public health policy.

The intensity of a dose of nature contact includes the size of the natural area or park a person visits (Bratman et al., 2019; Shanahan, Fuller, Bush, Lin, and Gaston, n.d.). Experimental approaches to nature contact are limited in the number of natural areas they can integrate into their study designs van den Bosch and Sang, 2017. A prior study found that the visitors to the largest parks in San Francisco exhibited the greatest mental benefits (Schwartz et al., 2019). Here, we hypothesize that larger parks will provide greater mental benefits to those who visit them in cities, in general.

Studies using data from mobile phone applications and Twitter have sampled over a time period between weeks and months and have not been able to verify whether the timing (e.g., hour of day, day of week, time of year) of park visits impacts potential health benefits. However, a study using tweets in Melbourne demonstrated heterogeneity in emotional responses to nature across different seasons and time of day (Lim et al., 2018). In addition, comparing the benefits of park visitation temporally is a way to check the extent to which observed happiness in parks is a function of park visits occurring during the weekend or summer vacation.

Here, we expand our prior work in San Francisco to the 25 largest cities in the US by population. For each city, we estimate a similar metric of *happiness benefit*. We compare this indicator across cities using data from a four year period. We also compare the happiness benefit across different categories of park size, as well as across levels of city-wide park investment and quality. Finally, we compare the happiness benefit of park visitation among different seasons and days of the week.

## 3.3 Methods

### 3.3.1 Data Collection & Processing

We used a database of tweets collected from January 1 2012 to April 27 2015 (Appendix 7.1.1), limiting our search to English language tweets that included GPS coordinate location data (latitude and longitude). We chose this time period because geo-located tweets became abundant nationally in 2012 and dropped significantly in April 2015 when Twitter made precise location sharing an opt-in feature. Using boundaries from the US Census, we collected tweets within each of the 25 largest cities in the US by population (U.S. Census Bureau, 2012). We did not include retweets (tweets that are re-posted from another user) in our analysis.

We detected whether a tweet was posted within park boundaries using the Trust for Public Land’s Park Serve database. Our ability to find tweets posted from inside parks depends on the accuracy of mobile GPS hardware which can vary by manufacturer, surrounding building height, and weather conditions. While most message locations should be precise to within 10m, some of our user pool may have posted just outside of parks due to measurement error. Data analysis of hashtag frequency revealed that a large number of geo-located tweets were posted by automated accounts (or bots) posting about job opportunities and traffic; any tweet found with a job or traffic related hashtag was removed from the sample (Appendix 7.1.3).

We assigned a control tweet to each in-park tweet. For each tweet, we chose the closest-in-time out-of-park tweet from another user, temporally proximate to the in-park tweet within the same city. This message functions as a control because it allows us to compare the happiness of our in-park sample with a set of tweets

that were posted in the same city and at roughly the same time. We summarize each city’s Twitter data in Table 3.1. In Appendix 7.1.4, we describe an alternative control group specification that uses out-of-park tweets from the same users who posted tweets inside of parks.

### 3.3.2 Sentiment Analysis

To understand the mental benefits of park visitation, we used sentiment analysis, a natural language processing technique that associates numerical values to the emotional response induced by individual words. For the present study, we used the Language Assessment by Mechanical Turk (labMT) sentiment dictionary which includes 10,222 of the most commonly used English words, merged from four distinct text corpora, and rated on a scale of 1 (least happy) to 9 (most happy) (Dodds, Harris, Kloumann, Bliss, and Danforth, 2011). For example, *beautiful* has an average happiness score of 7.92, *city* has an average happiness score of 5.76, and *garbage* has an average happiness score of 3.18 in labMT. We excluded words with scores between 4.0 and 6.0 from our analysis because they are emotionally neutral or particularly context dependent. The labMT sentiment dictionary performs well when compared with other sentiment dictionaries on large-scale texts, and correlates with traditional surveys of well-being including Gallup’s well-being index (Mitchell, Frank, Harris, Dodds, and Danforth, 2013; Reagan, Danforth, Tivnan, Williams, and Dodds, 2017). When using this type of bag-of-words approach, it is inappropriate to rate the happiness levels of individual tweets.

For each round of analysis, we aggregated tweets into an in-park group and a control group. We calculated the average happiness for each group of tweets as the weighted average of their labMT word scores using relative word frequencies as weights

Table 3.1: Summary of geolocated Twitter data for the 25 most populous cities in the US, collected 2012-2015. See Appendix 7.1.1 for details.

<u>City</u>	<u>Total</u> <u>tweets</u>	<u>Park</u> <u>tweets</u>	<u>% tweets</u> <u>in parks</u>	<u>Park</u> <u>visitors</u>	<u>Parks</u> <u>visited</u>	<u>Tweets</u> <u>per capita</u>
New York	2,892,512	213,813	7.4	113,702	1,880	0.35
Los Angeles	1,215,288	53,988	4.4	36,271	540	0.32
Philadelphia	1,166,125	64,857	5.6	26,287	482	0.76
Chicago	1,130,611	66,100	5.8	36,919	872	0.41
Houston	821,433	39,581	4.8	13,464	501	0.38
San Antonio	589,595	23,566	4.0	12,763	268	0.43
Washington	570,157	74,937	13.1	41,062	370	0.92
Boston	547,625	52,689	9.6	23,479	682	0.87
San Diego	491,219	36,080	7.3	22,269	406	0.37
Dallas	490,918	21,787	4.4	12,211	346	0.40
San Francisco	486,782	59,412	12.2	36,175	407	0.59
Austin	449,853	23,547	5.2	14,689	289	0.55
Baltimore	333,734	12,965	3.9	5,135	260	0.53
Fort Worth	320,178	9,664	3.0	4,278	239	0.42
Phoenix	268,455	12,041	4.5	7,566	189	0.18
Columbus	251,573	8,884	3.5	4,340	328	0.31
San Jose	234,234	8,263	3.5	4,517	314	0.24
Indianapolis	225,931	11,560	5.1	5,660	183	0.27
Charlotte	218,310	8,039	3.7	3,868	190	0.29
Seattle	201,533	12,758	6.3	7,739	373	0.32
Detroit	195,572	7,885	4.0	3,819	234	0.28
Jacksonville	194,777	6,219	3.2	3,218	261	0.23
Memphis	137,222	5,614	4.1	3,112	163	0.21
Denver	131,240	6,243	4.8	3,902	279	0.21
El Paso	96,015	2,722	2.8	1,397	180	0.14



in Equation 3.1:

$$h_{avg} = \frac{\sum_i^N h_i \cdot f_i}{\sum_i^N f_i}, \quad (3.1)$$

where  $h_i$  is the happiness score of the  $i$ th word and  $f_i$  is its frequency in a group of tweets with  $N$  words. Next, we subtracted the average happiness of the control tweets from the average happiness of the in-park tweets and defined this difference as the “happiness benefit”. To estimate uncertainty in our calculation of happiness benefit, we applied a bootstrapping procedure: We randomly sampled 80% of tweets without replacement from a set of in-park tweets and their respective control tweets and then re-calculated the happiness benefit. Performing this procedure 10 times, we derived a range of plausible happiness benefit values. Robustness checks were performed to show the convergence of this range at 10 runs.

We used the above technique to calculate the happiness benefit for all cities together and each city individually. For each city, we removed all words appearing in that city’s park name before estimating the happiness benefit. For example, we removed *golden*, with an average happiness of 7.3, from all San Francisco tweets because of Golden Gate Park. The word *park* is also removed from all tweets. We performed a manual check on the top ten most influential words in a city’s happiness benefit calculation. This allowed us to identify potential biases introduced by words being used in an unexpected manner. For example, we removed *ma* from all Boston tweets because it appears with a high frequency as an abbreviation for Massachusetts, but has a positive happiness score as shorthand for *mother*. We include the full list of stop words in Appendix 7.1.2.

### 3.3.3 Park Analysis

We used data from the Trust for Public Land (TPL) to further investigate the happiness benefit from urban park visits. The TPL provides a variety of data on municipal park systems. Annually, TPL publishes a ParkScore® for the largest cities in the US, which is a composite score out of 100 that combines metrics of park size, access, investment, and amenities. We conducted a correlation analysis for city-level happiness benefit against 2018 ParkScore® and park spending per capita, also sourced from the TPL (The Trust for Public Land, 2019). ParkScore® and spending for Indianapolis was sourced from TPL’s 2017 data release due to lack of participation in 2018.

To investigate the relationship between happiness benefit and park size, we assigned every in-park tweet a category based on the size of the park from where it was posted. We grouped parks into four categories ( $< 1$  acre, between 1 and 10 acres, between 10 and 100 acres, and greater than 100 acres). To have roughly equal representation from each city, we randomly selected tweets (along with their control tweet) in each park category from each city (or all of the tweets in that category if there were less than 500). After combining the randomly selected tweets from each city for each park category, we estimated the happiness benefit using the same bootstrapping procedure described above.

### 3.3.4 Temporal Analysis

Next, we estimated the happiness benefit based on when tweets were posted in three different ways. First, we grouped tweets based on the month they were posted in four seasonal groups (Winter: Dec, Jan, Feb; Spring: Mar, Apr, May; Summer: Jun, Jul, Aug; Fall: Sep, Oct, Nov). Second, we grouped tweets based on the day of the week

they were posted. Finally, we grouped tweets based on the hour of the day they were posted in their local timezone (Appendix 7.1.5). To have roughly equal representation from each city, we randomly selected 1,000 tweets (along with their control tweet) in each time category from each city (or all of the tweets in that category if there were less than 1,000). After combining the randomly selected tweets from each city, we estimated the happiness benefit using the same bootstrapping procedure described above.

## **3.4 Results**

### **3.4.1 Sentiment Analysis**

Across all cities, the mean happiness benefit was 0.10 (Bootstrap Range [.098, .103]). Across our 25 city sample, the mean happiness benefit ranged from 0.00 to 0.18. Indianapolis had the highest mean happiness benefit, while Baltimore had the lowest (Fig. 3.1). Cities with more in-park tweets to sample from had tighter happiness benefit ranges, as exhibited by Denver, New York, Los Angeles, and Philadelphia. The mean happiness benefit was positive across all cities.

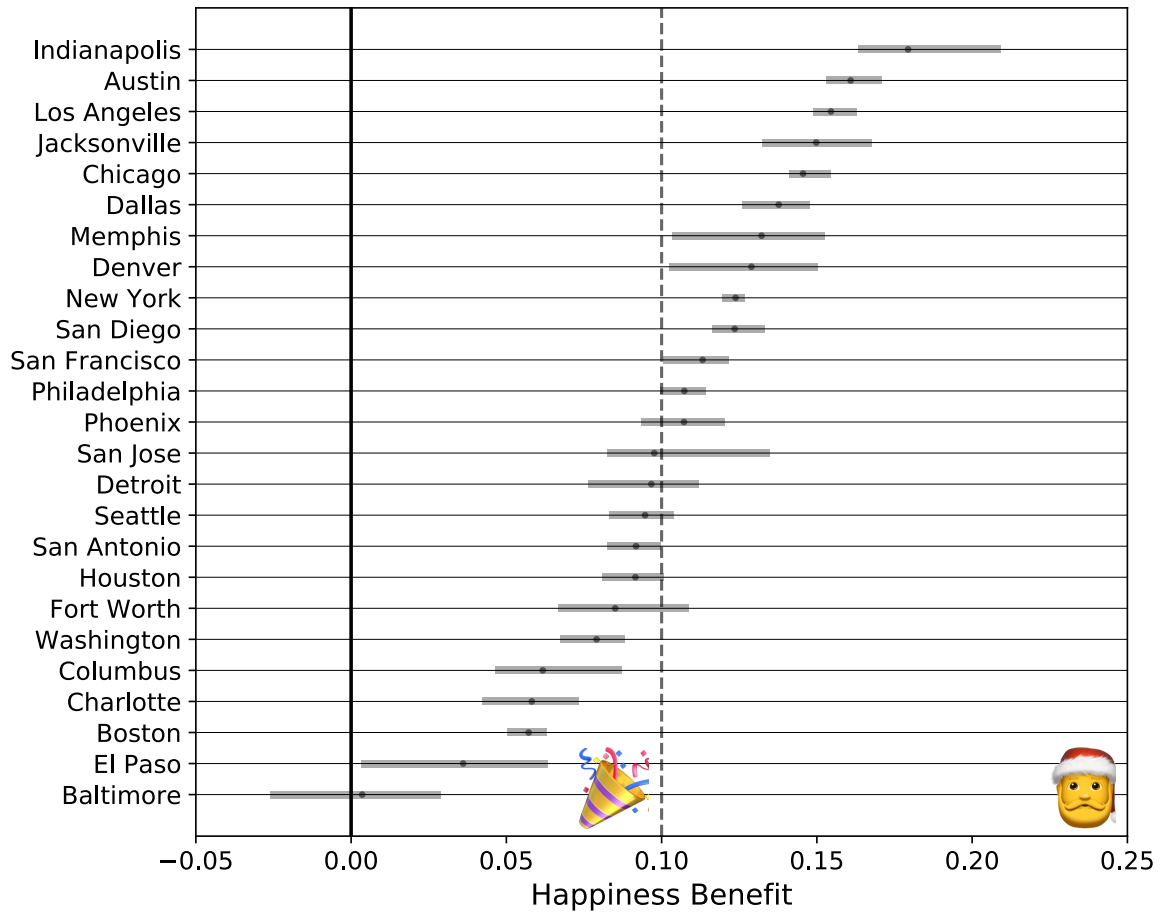


Figure 3.1: Happiness benefit by city, full range of values from 10 bootstrap runs in which 80% of tweets were randomly selected. The dark grey dots represent the mean value from bootstrap runs. For each city, the control group consists of non-park tweets posted at roughly the same time as each in-park tweet. The solid line marks a happiness benefit of 0, and the dotted line is average happiness benefit across all 25 cities. Emojis denote the happiness benefit typically observed on New Year's Day and Christmas for all English tweets.

## Wordshifts

The happiness benefit is driven by word frequency differences between the in-park tweets and control tweets. Specifically, positive words (with a happiness score greater than 6) including *beautiful*, *fun*, *enjoying*, and *amazing* appeared more frequently in in-parks tweets. Negative words (with a happiness score less than 4) such as *don't*, *not* and *hate* appeared less frequently in in-park tweets. We illustrate the variation in relative frequencies in Fig. 3.2, a wordshift plot that demonstrates the most influential words (by frequency and happiness) driving the happiness benefit (Dodds et al., 2011). Interactive versions of the city wordshift graphs are available in the online appendix accompanying this manuscript at <http://compstorylab.org/cityparkhappiness/>.

### 3.4.2 Park Analysis

We plot the mean happiness benefit values against two metrics of park quality — park spending and ParkScore® (Fig. 3.3). There is no clear pattern between happiness benefit and park spending or ParkScore®. Interestingly, Indianapolis, which had the highest mean happiness benefit, had the lowest municipal park spending per capita and one of the lowest ParkScore® values. Washington D.C., San Francisco, Chicago, New York, and Seattle had the highest ParkScore® values, and were all fairly close to the mean happiness benefit of 0.10.

We grouped in-park tweets into four categories based on the size of the park and estimated the happiness benefit for each category. Parks greater than 100 acres had the highest mean happiness benefit of 0.13, followed by parks from 1 – 10 acres (0.12). Parks less than 1 acre and parks between 10–100 acres had the lowest mean happiness benefit of 0.09 (Fig. 3.5).

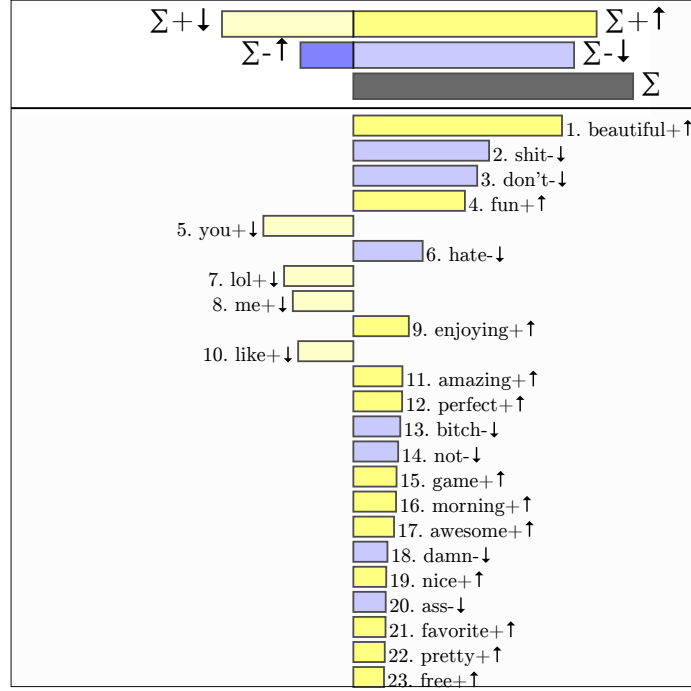


Figure 3.2: Differences in word frequency between park and control tweets across all cities, in order of decreasing contribution to the difference in average happiness. The right side represents the park tweets, with an average happiness of 5.96. The left side represents the control tweets, with an average happiness of 5.86. Purple bars represent words  $\leq 4$  (with  $-$  symbol) on the Hedonometer scale. Yellow bars represent words  $\geq 6$  (with  $+$  symbol) on the Hedonometer scale. Arrows indicate whether a word was more or less frequent within that set of tweets compared to the other text. For example, beautiful is a positive word (yellow) with higher frequency in in-park tweets that contributes to its higher average happiness than the control tweets. Don't is a negative word (purple) that appears less frequently in in-park tweets, also resulting in a higher average happiness score compared to control groups. Going against the overall trend, the positive words lol and me are used less often in parks. This wordshift uses tweets from 1,000 random in-park tweets and 1,000 control tweets from each city.

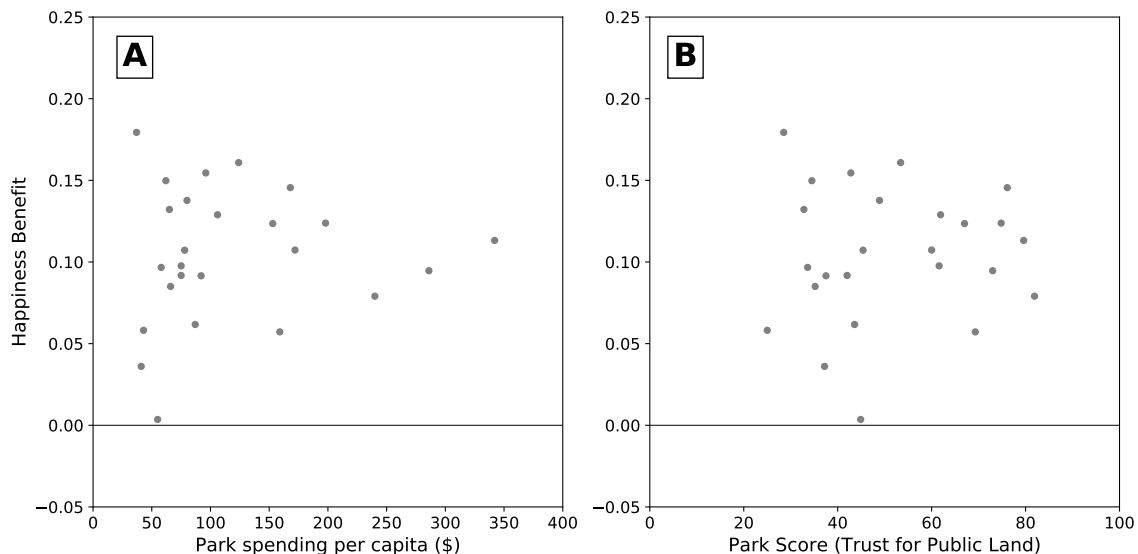


Figure 3.3: A. The left panel shows park spending per capita vs mean happiness benefit by city. Park spending per capita is from Trust for Public Land (TPL) data. B. The right panel shows ParkScore® vs happiness benefit. The TPL calculates ParkScore® annually from measures of park acreage, access, investment, and amenities, and is scaled to a maximum score of 100. The happiness benefit was not strongly correlated with per capita spending (Spearman’s  $\rho = 0.14$ ) or ParkScore® (Spearman’s  $\rho = 0.03$ ).

## Comparing Cities

We analyzed the average happiness of each individual city’s tweets. For example, Chicago had over 1.1 million total tweets, with 36,919 users tweeting from a park. Tweets were posted in 872 separate park units, second only to New York. Roughly 6% of all Chicago tweets were posted from a park, with .41 tweets per capita from 2012–2015. Chicago’s happiness benefit was 0.15, ranking fifth among our 25 cities. Chicago’s tweets were distributed among many different types of parks, including several large parks along the shore of Lake Michigan. Tweets posted in Chicago parks had higher average happiness than tweets posted elsewhere in Chicago due to higher frequency of happy words such *beautiful* and *great*, and lower frequency of unhappy words including profanity, *don’t*, and *not*. We include a map of Chicago’s parks and a wordshift plot between Chicago’s in-park and control tweets in Fig. 3.4.

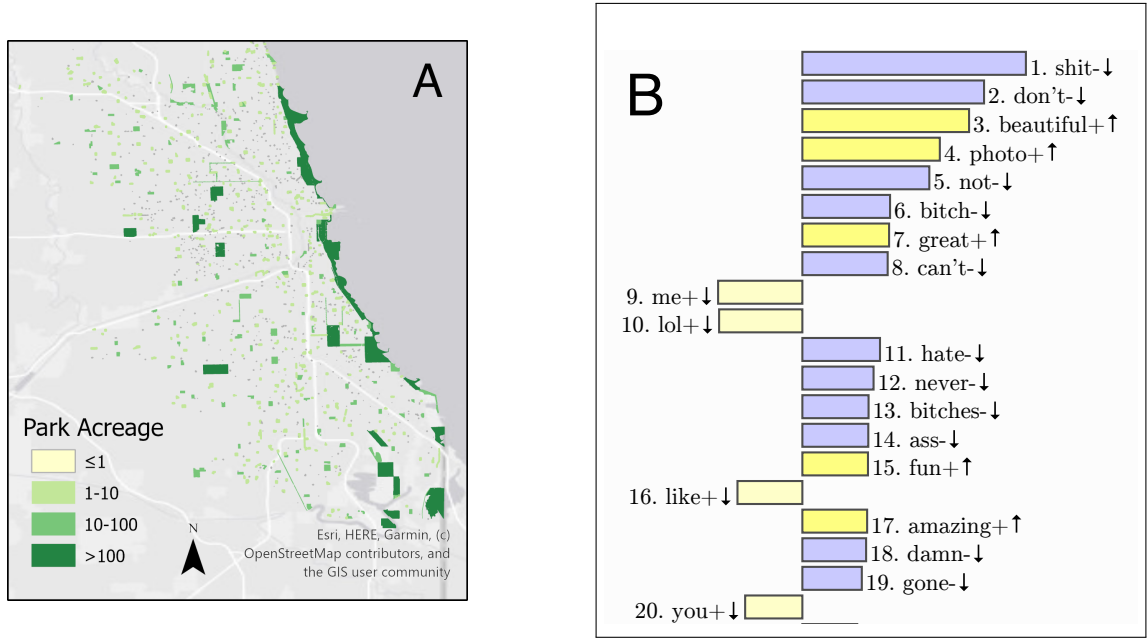


Figure 3.4: A. The left panel shows a map of the greater Chicago area and its municipal parks, shaded by park size category. B. The right panel is a wordshift for Chicago's tweets. In this wordshift the right side represents the park tweets. The left side represents the control tweets. Purple bars represent words  $\leq 4$  (with  $-$  symbol) on the Hedonometer scale. Yellow bars represent words  $\geq 6$  (with  $+$  symbol) on the Hedonometer scale. Arrows indicate whether a word was more or less frequent within that set of tweets compared to the other text. Individuals use positive words such as beautiful and fun more often in Chicago parks, and use profanity less often.



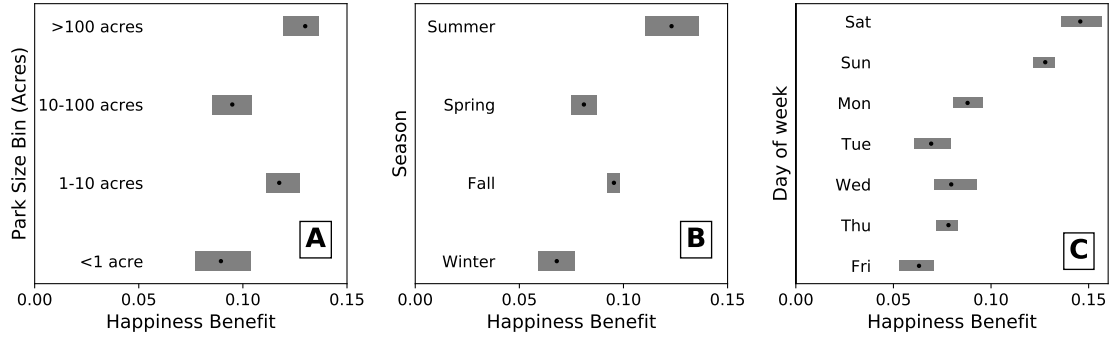


Figure 3.5: A. The left panel shows happiness benefit by park size. The largest category of parks (greater than 100 acres) had the highest happiness benefit. B. The middle panel shows happiness benefit by season, with summer and fall exhibiting the highest mean happiness benefit values. C. The right panel shows happiness benefit by day of the week, with the weekend days higher than other days of the week. In all three panels, the range is the full range of happiness benefits from 10 runs, sampling 80% of tweets. 1,000 random in-park tweets were pooled in each group from each city. Control tweets were selected as tweets most temporally proximate to the in-park tweet from the same city.

### 3.4.3 Temporal Analysis

Across all cities, we grouped park tweets and their control tweets according to the in-park tweet’s timestamp. First, we compared the happiness benefit by season. The mean happiness benefit was highest in the summer (0.12), followed by fall (0.10), spring (0.08), and winter (0.06) as shown in Fig. 3.5. Then we grouped park tweets and their respective control tweets according to the day of the week in which it was posted. Saturday exhibited the highest mean happiness benefit (.15) followed by Sunday (0.13). Monday through Friday were all between 0.06 and 0.09 (Fig. 3.5). We also estimated the happiness benefit by hour of the day in which the in-park tweet was posted. The tweets posted during the 8:00 and 9:00 AM hours had a mean happiness benefit around 0.07 while the rest of the day did not show a clear pattern, ranging from 0.08 to 0.14 (Appendix 7.1.5).

## 3.5 Discussion

### 3.5.1 Sentiment Analysis

In this study, across the 25 largest cities in the US, we find that people write happier words on Twitter in parks than they do outside of parks. This effect is strongest for the largest parks by area - greater than 100 acres. The effect is present during all seasons and days of the week, but is most prominent during the summer and on weekend days.

Pooling tweets across cities, we find a mean happiness benefit of 0.10. According to Hedonometer.org, which tracks Twitter happiness as a whole using the labMT dictionary, Twitter has fluctuated around a mean happiness of 6.02 since 2008. New Year's Day has historically had an average happiness of 6.11, giving it an average happiness benefit of .10. Christmas, historically the happiest day of the year on Twitter, has had an average happiness benefit of 0.24. The global COVID-19 Pandemic gained rapid recognition in the US on March 12, 2020, which resulted in the then unhappiest day in Twitter's history with a drop of 0.31 from its historical average. Following the murder of George Floyd, the Black Lives Matter protests led to a new all-time low, 0.39 below the historical average (Hedonometer.org, 2020). These are considered large swings, and we assess that the happiness benefit of 0.10 across a sample of 25,000 tweets is a strong signal.

Positive words such as *beautiful*, *fun*, and *enjoying* contributed to the higher levels of happiness from our in-park tweet group. These words may relate to the stimulating aspects of urban greenspace. This is supported by a recent study that analyzed tweets to investigate which aspects of restoration were most prominent in urban greenspace. They found that fascination, an emotional state induced

through inherently interesting stimuli, was most salient (Wilkie, Thompson, Cranner, and Ginty, 2020). Fascination is one characteristic of nature experiences described by Attention Restoration Theory, which theorizes that time in nature provides an opportunity to recover from the cognitive fatigue induced by mentally taxing urban environments (R. Kaplan and Kaplan, 1989; S. Kaplan, 1995).

We find high levels of variation across cities for the happiness benefit between in-park and out-of-park tweets. In Chicago, higher frequencies of words such as *beautiful* drive higher in-park tweet happiness. Park tweets had lower frequencies of negative words such as *don't*, *not*, and *hate* (Fig. 3.4). Psychological experiments treat positive and negative affect as separate measures (McMahan and Estes, 2015); the heterogeneity of the words driving the happiness benefit may be related to how these components of affect are being expressed via tweets.

### 3.5.2 Park Analysis

Park spending per capita and ParkScore® were not correlated with mean happiness benefit by city. However, prior work has demonstrated an association between park investment and levels of self-rated health (Mueller et al., 2019). Another study found higher levels of physical activity and health to be associated with a composite score of park quality in 59 cities (Mullenbach, Mowen, and Baker, 2018). Other factors such as heterogeneous use patterns of Twitter across cities may be more associated with happiness benefit than measures of park quality and spending. We encourage further investigation of the relationship between park quality and investment with the mental health benefits of nature contact.

Tweets inside of all park size categories exhibited a positive happiness benefit. The largest parks, greater than 100 acres, had the highest mean happiness benefit. One possible explanation is that larger parks provide greater opportunities for mental

restoration and separation from the taxing environment of the city. This finding is consistent with results from our earlier study in San Francisco, in which tweets in the larger and greener Regional Parks had the highest happiness benefit (Schwartz et al., 2019). Parks between 0 and 10 acres are often neighborhood parks that people use in their day to day lives. Local parks provide many essential functions; however, our results suggest that the experiences people have in larger parks may be more beneficial from a mental health perspective. Another possibility is that people spend more time in larger parks; one study suggested that 120 minutes of nature contact a week resulted in improved health and well-being (White et al., 2019).

### **3.5.3 Temporal Analysis**

We observe that the mean happiness benefit was higher in summer than other seasons; however, the happiness benefit was positive in all four seasons. Similarly, the mean happiness benefit was highest during the weekend, but positive on all days of the week (Fig. 3.5). People use happier words when visiting parks throughout the week and year — not just outside of typical working hours. This result is encouraging because some prior studies on nature contact using Twitter only addressed shorter time spans. Future studies should seek methods that can investigate the other temporal aspects of nature contact including the frequency and duration of visits (Shanahan et al., n.d.).

### **3.5.4 Future Directions**

Future research should continue to explore the relationship between tweet happiness and other factors beyond park investment. While ParkScore® captures a variety of park-quality related metrics, vegetation and biodiversity are salient features of

greenspace that significantly impact how people experience their time in nature (Clark et al., 2014; Mavoa, Davern, Breed, and Hahs, 2019; Wang, Kotze, Vierikko, and Niemelä, 2019). More localized studies could look at the mental health impact of park-level vegetative cover and biodiversity metrics. While we investigated the seasonal variation of in-park happiness, climate and weather have been shown to influence happiness on Twitter as well (Baylis et al., 2018; Moore, Obradovich, Lehner, and Baylis, 2019). Tweets could be binned by some composite of temperature, humidity, and precipitation in order to investigate how weather moderates the association between nature contact and mental well-being (van den Bosch and Sang, 2017). Some greenspaces are more crime prone than others and a recent study was able to identify crime-related tweets, which may help further explain happiness differences between parks (Curiel, Cresci, Muntean, and Bishop, 2020; Kimpton, Corcoran, and Wickes, 2017). Demographic, socioeconomic, and cultural factors also play a role in how people engage with parks (Browning and Rigolon, 2018). While identifying such factors on Twitter is challenging and requires ethical consideration, other methodologies can continue to explore how different groups use and benefit from time in parks, to help ensure that the benefits of parks are available to everyone. As the evidence continues to mount on the many different benefits of nature contact, ensuring access to quality parks for all urban residents is critical.

## 3.6 References

- Baylis, P., Obradovich, N., Kryvasheyeu, Y., Chen, H., Coviello, L., Moro, E., ... Fowler, J. H. (2018). Weather impacts expressed sentiment. *PloS one*, 13(4). DOI: [10.1371/journal.pone.0195750](https://doi.org/10.1371/journal.pone.0195750)

- Berto, R. (2014). The Role of Nature in Coping with Psycho-Physiological Stress: A Literature Review on Restorativeness. *Behavioral Sciences*, 4(4), 394–409.  
DOI: [10.3390/bs4040394](https://doi.org/10.3390/bs4040394)
- Bratman, G., Anderson, C. B., Berman, M. G., Cochran, B., De Vries, S., Flanders, J., ... Hartig, T., et al. (2019). Nature and mental health: An ecosystem service perspective. *Science advances*, 5(7), eaax0903.
- Browning, M. H., & Rigolon, A. (2018). Do income, race and ethnicity, and sprawl influence the greenspace-human health link in city-level analyses? findings from 496 cities in the united states. *International journal of environmental research and public health*, 15(7), 1541.  
DOI: [10.3390/ijerph15071541](https://doi.org/10.3390/ijerph15071541)
- Clark, N. E., Lovell, R., Wheeler, B. W., Higgins, S. L., Depledge, M. H., & Norris, K. (2014). Biodiversity, cultural pathways, and human health: A framework. *Trends in ecology & evolution*, 29(4), 198–204.  
DOI: [10.1016/j.j.tree.2014.01.009](https://doi.org/10.1016/j.j.tree.2014.01.009)
- Curiel, R. P., Cresci, S., Muntean, C. I., & Bishop, S. R. (2020). Crime and its fear in social media. *Palgrave Communications*, 6(1), 1–12.  
DOI: [10.1057/s41599-020-0430-7](https://doi.org/10.1057/s41599-020-0430-7)
- Dodds, P. S., Harris, K. D., Kloumann, I. M., Bliss, C. A., & Danforth, C. M. (2011). Temporal patterns of happiness and information in a global social network: Hedonometrics and twitter. *PloS one*, 6(12).  
DOI: [10.1371/journal.pone.0026752](https://doi.org/10.1371/journal.pone.0026752)
- Frumkin, H., Bratman, G. N., Breslow, S. J., Cochran, B., Kahn Jr, P. H., Lawler, J. J., ... Wood, S. A. (2017). Nature contact and human health: A research agenda. *Environmental Health Perspectives*, 125(7), 1–18.  
DOI: [10.1289/EHP1663](https://doi.org/10.1289/EHP1663)
- Hedonometer.org. (2020). Hedonometer, from The Vermont Complex Systems Center & Computational Story Lab. <http://hedonometer.org/>.
- Houlden, V., Weich, S., de Albuquerque, J. P., Jarvis, S., & Rees, K. (2018). The relationship between greenspace and the mental wellbeing of adults: A systematic review. *PloS one*, 13(9).  
DOI: [10.1371/journal.pone.0203000](https://doi.org/10.1371/journal.pone.0203000)
- Kaplan, R., & Kaplan, S. (1989). *The experience of nature: A psychological perspective*. CUP Archive.
- Kaplan, S. (1995). The restorative benefits of nature: Toward an integrative framework. *Journal of Environmental Psychology*, 15(3), 169–182.  
DOI: [10.1016/0272-4944\(95\)90001-2](https://doi.org/10.1016/0272-4944(95)90001-2)

- Kimpton, A., Corcoran, J., & Wickes, R. (2017). Greenspace and crime: An analysis of greenspace types, neighboring composition, and the temporal dimensions of crime. *Journal of Research in Crime and Delinquency*, 54(3), 303–337.  
DOI: [10.1177/0022427816666309](https://doi.org/10.1177/0022427816666309)
- Krabbendam, L., van Vugt, M., Conus, P., Söderström, O., Empson, L. A., van Os, J., & Fett, A.-K. J. (2020). Understanding urbanicity: How interdisciplinary methods help to unravel the effects of the city on mental health. *Psychological Medicine*, 1–12.  
DOI: [10.1017/S0033291720000355](https://doi.org/10.1017/S0033291720000355)
- Lim, K. H., Lee, K. E., Kendal, D., Rashidi, L., Naghizade, E., Winter, S., & Vasardani, M. (2018). The grass is greener on the other side: Understanding the effects of green spaces on Twitter user sentiments. In *Companion of the the web conference 2018 on the web conference 2018 - www '18* (pp. 275–282).  
DOI: [10.1145/3184558.3186337](https://doi.org/10.1145/3184558.3186337)
- Mavoa, S., Davern, M., Breed, M., & Hahs, A. (2019). Higher levels of greenness and biodiversity associate with greater subjective wellbeing in adults living in melbourne, australia. *Health & place*, 57, 321–329.  
DOI: [10.1016/j.healthplace.2019.05.006](https://doi.org/10.1016/j.healthplace.2019.05.006)
- McDonald, R. I., Beatley, T., & Elmquist, T. (2018). The green soul of the concrete jungle: the urban century, the urban psychological penalty, and the role of nature. *Sustainable Earth*, 1x'1(1), 3.  
DOI: [10.1186/s42055-018-0002-5](https://doi.org/10.1186/s42055-018-0002-5)
- McMahan, E. A., & Estes, D. (2015). The effect of contact with natural environments on positive and negative affect: A meta-analysis. *The Journal of Positive Psychology*, 10(6), 507–519.  
DOI: [10.1080/17439760.2014.994224](https://doi.org/10.1080/17439760.2014.994224)
- Mitchell, L., Frank, M. R., Harris, K. D., Dodds, P. S., & Danforth, C. M. (2013). The geography of happiness: Connecting twitter sentiment and expression, demographics, and objective characteristics of place. *PLoS ONE*, 8(5), e64417.  
DOI: [10.1371/journal.pone.0064417](https://doi.org/10.1371/journal.pone.0064417)
- Moore, F. C., Obradovich, N., Lehner, F., & Baylis, P. (2019). Rapidly declining remarkability of temperature anomalies may obscure public perception of climate change. *Proceedings of the National Academy of Sciences*, 116(11), 4905–4910.  
DOI: [10.1073/pnas.1816541116](https://doi.org/10.1073/pnas.1816541116)
- Mueller, J. T., Park, S. Y., & Mowen, A. J. (2019). The relationship between self-rated health and local government spending on parks and recreation in the

- united states from 1997 to 2012. *Preventive medicine reports*, 13, 105–112.  
DOI: [10.1016/j.pmedr.2018.11.018](https://doi.org/10.1016/j.pmedr.2018.11.018)
- Mullenbach, L. E., Mowen, A. J., & Baker, B. L. (2018). Peer reviewed: Assessing the relationship between a composite score of urban park quality and health. *Preventing chronic disease*, 15.  
DOI: [10.5888/pcd15.180033](https://doi.org/10.5888/pcd15.180033)
- Reagan, A. J., Danforth, C. M., Tivnan, B., Williams, J. R., & Dodds, P. S. (2017). Sentiment analysis methods for understanding large-scale texts: a case for using continuum-scored words and word shift graphs. *EPJ Data Science*, 6(1), 28.  
DOI: [10.1140/epjds/s13688-017-0121-9](https://doi.org/10.1140/epjds/s13688-017-0121-9)
- Rigolon, A., Browning, M., & Jennings, V. (2018). Inequities in the quality of urban park systems: An environmental justice investigation of cities in the united states. *Landscape and Urban Planning*, 178, 156–169.  
DOI: [10.1016/j.landurbplan.2018.05.026](https://doi.org/10.1016/j.landurbplan.2018.05.026)
- Schwartz, A. J., Dodds, P. S., O’Neil-Dunne, J. P., Danforth, C. M., & Ricketts, T. H. (2019). Visitors to urban greenspace have higher sentiment and lower negativity on twitter. *People and Nature*, 1(4), 476–485.  
DOI: [10.1002/pan.3.10045](https://doi.org/10.1002/pan.3.10045)
- Shanahan, D. F., Fuller, R. A., Bush, R., Lin, B. B., & Gaston, K. J. (n.d.). The health benefits of urban nature: how much do we need? *BioScience*, (5), 476–485.  
DOI: [10.1093/biosci/biv032](https://doi.org/10.1093/biosci/biv032)
- Taylor, L., Hahs, A. K., & Hochuli, D. F. (2018). Wellbeing and urban living: Nurtured by nature. *Urban Ecosystems*, 21(1), 197–208.  
DOI: [10.1007/s11252-017-0702-1](https://doi.org/10.1007/s11252-017-0702-1)
- The Trust for Public Land. (2019). Park serve. data retrieved from TPL, <https://www.tpl.org/parkserve/downloads>.
- U.S. Census Bureau. (2012). U.S. Census Populated Places. data retrieved from ESRI, <https://hub.arcgis.com/datasets/esri::usa-census-populated-places>.
- Van den Berg, M., Wendel-Vos, W., van Poppel, M., Kemper, H., van Mechelen, W., & Maas, J. (2015). Health benefits of green spaces in the living environment: A systematic review of epidemiological studies. *Urban Forestry & Urban Greening*, 14(4), 806–816.  
DOI: [10.1016/j.ufug.2015.07.008](https://doi.org/10.1016/j.ufug.2015.07.008)
- van den Bosch, M., & Sang, Å. O. (2017). Urban natural environments as nature-based solutions for improved public health—a systematic review of reviews.



*Environmental research*, 158, 373–384.

DOI: [10.1016/j.envres.2017.05.040](https://doi.org/10.1016/j.envres.2017.05.040)

Wang, Y., Kotze, D. J., Vierikko, K., & Niemelä, J. (2019). What makes urban greenspace unique—relationships between citizens' perceptions on unique urban nature, biodiversity and environmental factors. *Urban Forestry & Urban Greening*, 42, 1–9.

DOI: [10.1016/j.ufug.2019.04.005](https://doi.org/10.1016/j.ufug.2019.04.005)

White, M. P., Alcock, I., Grellier, J., Wheeler, B. W., Hartig, T., Warber, S. L., ... Fleming, L. E. (2019). Spending at least 120 minutes a week in nature is associated with good health and wellbeing. *Scientific Reports*, 9(1), 7730.

DOI: [10.1038/s41598-019-44097-3](https://doi.org/10.1038/s41598-019-44097-3)

Wilkie, S., Thompson, E., Cranner, P., & Ginty, K. (2020). Attention restoration theory as a framework for analysis of tweets about urban green space: A case study. *Landscape Research*, 1–12.

DOI: [10.1080/01426397.2020.1738363](https://doi.org/10.1080/01426397.2020.1738363)

# Chapter 4: The health benefits of urban parks vary regionally across cities

## 4.1 Abstract

The health benefits of nature contact are receiving increased attention. However, there have been few national studies looking at the association between park access and health outcomes at the city level. In this study, we synthesized several publicly available data sets to investigate relationships between human health and park access, as well as the equity of park access itself among social and economic groups. We found park access to be associated with both improved mental health and reduced obesity across cities in the US. The associations between park access and both health outcomes were significant in the South and West, but not in the Northeast and Midwest regions of the US. Next, we examined the relationships between park access and socioeconomic and demographic factors at the census tract level. On average, park access was highest at the low and high ends of the income distribution. For race, we found that park access generally declines with White population and increases with Black population. These relationships suggest that the potential health benefits of nature contact may be obscured by complex inequities in opportunities to access parks.

## 4.2 Introduction

The COVID-19 pandemic and a national wave of protests following the murder of George Floyd have reminded us of the inequalities that permeate our urban areas as well as the importance of public spaces. Public parks are essential infrastructure that

provide residents with opportunities for exercise, socialization, and a refuge from the stress of urban life. Through these and other pathways, parks can promote a variety of health outcomes. However, few studies have compared park access and health outcomes across the entire US. In addition, park access varies within and across cities, which may result in unequal opportunities for people to take advantage of health benefits.

A general consensus has emerged that nature contact is associated with improved physical and mental well-being (Frumkin et al., 2017). A review of epidemiological studies found strong evidence for positive associations between quantity of greenspace and perceived mental health (Van den Berg et al., 2015). Several studies have found positive associations between greenspace and lower obesity prevalence (Kuo, 2015). Obesity and declining mental health have received particular attention in the nature contact literature due to the growing burden they are having on society.

However, a majority of urban greenspace studies have been at the neighborhood (sub-city) level. Cities are also an important unit of study because they represent discrete policy units with a specific cultural, socioeconomic, and geographic context. Cities also compete for competitive grants to fund park development (Rigolon, Browning, and Jennings, 2018). For these reasons, some studies have compared park systems and their association with health outcomes across cities. Park quantity was correlated with self-reported well-being across 44 major cities (and less strongly correlated with park access and quality) (Larson, Jennings, and Cloutier, 2016). In another study, park density was associated with higher levels of physical activity and a lower probability of being overweight in 85 cities (West, Shores, and Mudd, 2012). Comparing park access across multiple cities has been limited due to a lack of data; a national database has not been widely accessible until recently (Mullenbach, Mowen, and Baker, 2018).

Population health studies have demonstrated regional variation in health outcomes across the US (Moriarty, Zack, Holt, Chapman, and Safran, 2009). There is also evidence that the relationship between urban greenspace and health will vary regionally due to factors such as urban design, climate, and culture. For example, a study that compared nearby vegetation and obesity found opposing relationships for Phoenix, Arizona (greenspace was protective) and Portland, Oregon (greenspace was harmful) (Tsai, Davis, and Jackson, 2019). A study investigating urban vegetation and health across US cities found that these relationships were moderated by race and ethnicity, and suggested that regional differences could be obscuring the greenspace-health link (Browning and Rigolon, 2018). These results suggest that regional differences need to be taken into account if we are to establish the circumstances under which greenspace delivers health benefits.

If park access can promote health at the city level, equitable park access would ensure that these benefits are widely available. Prior work has indicated that park access is not evenly distributed among different socioeconomic and demographic groups at the neighborhood level (Wen, Zhang, Harris, Holt, and Croft, 2013). However, the relationship between park access and these factors is not straightforward. In a study of 10 cities, income was positively correlated with park access in 4 cities, negatively correlated with park access in 3, and had no significant correlation in the remaining 3 three cities (Nesbitt, Meitner, Girling, Sheppard, and Lu, 2019). We need a better understanding of how park access varies across socioeconomic and demographic gradients within and across cities in order to gauge where and for whom parks can promote better health outcomes.

In this study, we first analyze the association between park access and health outcomes at the city level and test whether this relationship varies regionally using a national database of local parks. We then examine how park access itself varies

at the census-tract level along socioeconomic and demographic factors. We use this analysis to answer three questions: (1) Is access to urban parks associated with improved mental health and decreased obesity across cities in the US? (2) Does this association vary regionally? (3) How equitably is park access distributed among social and economic groups at the neighborhood level?

## 4.3 Methods

### 4.3.1 Health Outcomes

The Center for Disease Control’s (CDC) 500 Cities data provide city and census-tract level estimates for a variety of health outcomes and related risk factors in the largest cities in the US (CDC, 2018a). We included 498 of the cities in our study, leaving out Honolulu, HI and Anchorage, AK. These 498 cities fall into 4 high-level regions as specified by the US Census Bureau’s Region classifications, which classifies the 48 lower states into Northeast, Midwest, South, and West.

The CDC 500 Cities estimates are based on surveys conducted by the Behavioral Risk Factor Surveillance System (BRFSS) (CDC, 2018b). The CDC estimates obesity based on self-reported height and weight values from the BRFSS survey. For mental health, the BRFSS asks, *Now thinking about your mental health, which includes stress, depression, and problems with emotions, for how many days during the past 30 days was your mental health not good?* (CDC, 2018b). Respondents who report  $\geq 14$  days are recorded as having *frequent mental distress* (FMD). Obesity and FMD status are recorded as binary response variables for individual survey respondents. Based on these surveys, the CDC estimates population level prevalence values for cities and their constituent census tracts. Therefore, The 500 Cities data represent synthetic

rather than actual prevalence of health outcomes (Kong and Zhang, 2020). Due to smaller sample sizes in the underlying data, census tract estimates are not as reliable as the city-level estimates. For the purposes of this analysis, we assumed the validity of 500 Cities estimates at the city level. We describe this data in further detail in Appendix 7.2.1.

### 4.3.2 Park Access Estimates

Compared to urban vegetation generally, public parks may be especially relevant for health outcomes for several reasons. There is evidence that actively managed spaces provide a higher positive impact on health compared to unstructured natural areas (Fan, Das, and Chen, 2011). Urban residents specifically use parks to exercise, socialize, and relax—actions which are considered part of the mechanistic pathways from nature contact to impacts on health (Kabisch, Qureshi, and Haase, 2015; Markevych et al., 2017). Parks can also be improved through design, maintenance, and programming (Larson et al., 2016). These factors make parks a promising lever for promoting health in cities.

To estimate access to urban parks, we used ParkServe<sup>®</sup>, a spatial data set compiled by the Trust for Public Land (TPL). The ParkServe<sup>®</sup> project collected municipal park maps for roughly 14,000 cities across the US. For each city park, the TPL included a corresponding polygon delineating the area within a 10-minute walk of that park based on a road network map provided by ESRI (The Trust for Public Land, 2019). We estimated a census tract’s park access by summing the geographic area of that tract within 10-minutes of at least one park and dividing by that tract’s total area. At the city level, we estimated total park access via a population weighted average of census tract park access using Equation 4.1:

$$CityAccess_i = \sum_{j=1}^n Access_j * \frac{Pop_j}{Pop_i} \quad (4.1)$$

$Access_j$  is the % area of census tract  $j$  within a 10 minute walk of a park.

$Pop_j$  is that tract's population.  $Pop_i$  is the city's total population.

### 4.3.3 Socioeconomic & Demographic Variables

Population health outcomes are influenced by socioeconomic, demographic, and environmental factors (Zhang, Tan, and Diehl, 2017). The CDC uses correlations between the BRFSS survey data and factors such as age, race/ethnicity, sex, education, and poverty status to estimate population-level prevalence of health outcomes and behaviors. We obtained these variables at the census-tract level from the American Community Survey 5-year estimates (U.S. Census Bureau, 2017). A summary of the data sources is presented in Table 4.1.

### 4.3.4 City Analysis: Park Access & Health

First, we explored the relationship between park access and health outcome at the city level. We estimated ordinary least squares (OLS) regression models with health outcomes (obesity and FMD) as the dependent variables and park access as the independent variable for both the full national sample and regional sub-samples. We also estimated a model using regions as fixed effects alongside park access for the national sample. We include region in these models to account for potential geographic variation in the relationship between park access and health outcomes (Tsai et al., 2018).

*Table 4.1: CDC: Center for Disease Control 500 Cities Data (2018). ACS: American Community Survey (2013-2017 5-year estimates) TPL: The Trust for Public Land ParkServe® (2018)*

<u>Variable</u>	<u>Description</u>	<u>Source</u>
Frequent Mental Distress	Percent of Population	CDC
Obesity	Percent of Population	CDC
Region	Geographic Region	US Census
Park Access (Tract)	% Area within 10-minute walk of park	TPL
Park Access (City)	% Area within 10-minute walk weighted by census tract population	TPL
Population	Total Population	ACS
Income	Median income	ACS
Race/Ethnicity	Population White (%)	ACS
Race/Ethnicity	Population Black (%)	ACS
Age	Median Age	ACS
Education	Has Bachelor's (%)	ACS



### 4.3.5 Tract Level Park Access & Equity

Following our city level analysis of health, we looked more closely at how park access varies at the neighborhood level in our sample of cities. Prior work has found that park access, contrary to expectations, does not always follow a linear pattern along socioeconomic and demographic/ethnic gradients (Rigolon et al., 2018; Wen et al., 2013). To better understand these relationships, we compared park access at the census-tract level across regions. We used Locally Weighted Scatterplot Smoothing (LOWESS) to visualize the relationships between socioeconomic and demographic factors with census-tract level park access (Jacoby, 2000). LOWESS is a non-parametric regression method for curve fitting and visualizing non-linear relationships. With over 26,000 census tracts in our sample, it is not possible to detect trends when plotting all points. LOWESS is useful in this case because it allows us to detect overall trend lines and non-linear relationships. LOWESS estimates were conducted with Python package `statsmodels` (Seabold and Perktold, 2010).

## 4.4 Results

The 498 cities in our sample varied substantially. Population ranged from roughly 43,000 (Burlington, VT) to 8.5 million (New York, NY). The cities were distributed among the four census geographic regions with 54 in the Northeast, 93 in the Midwest, 158 in the South, and 193 in the West. Across the 498 cities in our sample, the estimated mean obesity prevalence was 29.5%, while the prevalence of frequent mental distress (FMD) was 12.8% (Table 4.2). Population weighted park access varied across cities between 4.45% (Hoover, Alabama) and 99% (Somerville, Massachusetts). Cities in the Northeast had the highest mean park access at 78.1% while Southern cities had a mean park access of 43.6%.

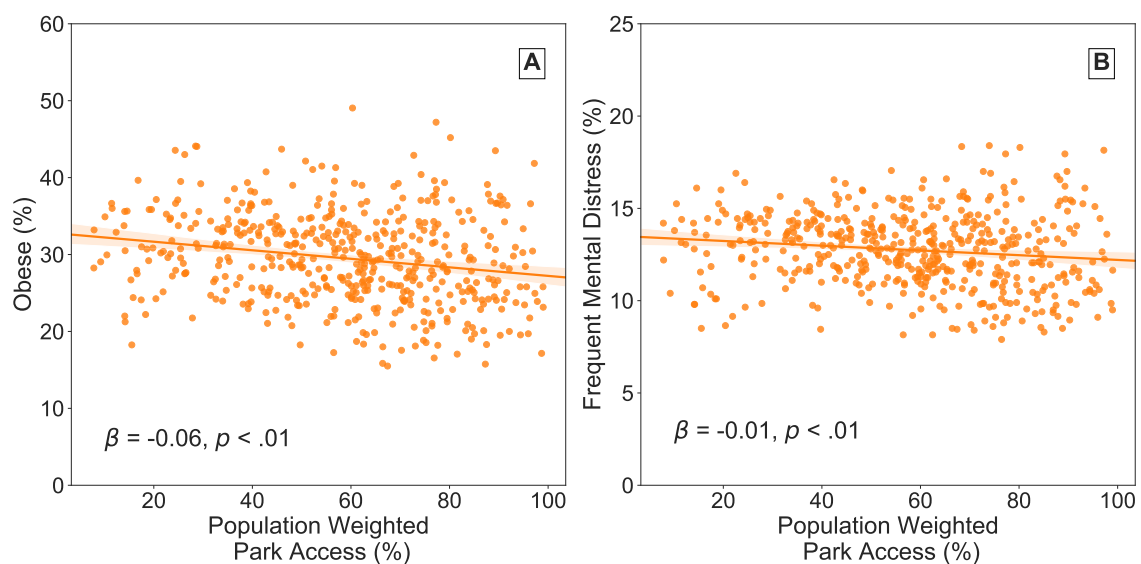
*Table 4.2: City counts and variable values by region. Park access is weighted based on census tract population. Values are means except for population, income, and age, which are medians within that group of cities.*

	All Cities	Midwest	Northeast	South	West
Obese (%)	29.53	32.62	31.01	32.04	25.58
Mental Health (%)	12.74	12.75	14.22	13.04	12.08
Park Access (%)	58.38	63.23	78.16	43.56	62.64
Population	112,283	101,928	95,143	131,151	110,153
White (%)	49.35	63.50	44.81	46.50	46.13
Black (%)	15.40	16.72	19.46	25.33	5.49
Income (\$)	27,085	26,908	24,211	26,005	30,008
Age	35	34.9	34.3	34.8	35.4
Has Bachelor's (%)	20.14	20.79	16.72	20.56	20.45

### 4.4.1 City Analysis: Park Access & Health

As population-weighted park access increases, obesity and FMD both trend downward (Fig. 4.1). The regression coefficients were weak but significant (Table 4.3, models 1 & 3). The regression models demonstrated that, on average, cities with higher park access had better health outcomes.

The relationship between park access and health outcomes varied across geographic regions. For obesity, park access regression coefficients were negative and significant in the South and West. (Fig. 4.2A-D). Similarly, for FMD, park access regression coefficients were negative and significant in the South and West (Fig. 4.2E-H).



*Figure 4.1: Scatterplots and regression lines for park access and health outcomes. Each point represents a city. Panel A: Park access and estimated obesity prevalence. Panel B: Park access and estimated FMD prevalence. Inset text displays the OLS regression coefficient  $\beta$  and p-value for population-weighted park access.*

Finally, we performed a multivariate OLS for all cities, including region as a categorical fixed effect in the model. Coefficients on park access remained small, but significant and negative. For obesity, the South and West regions had significant fixed

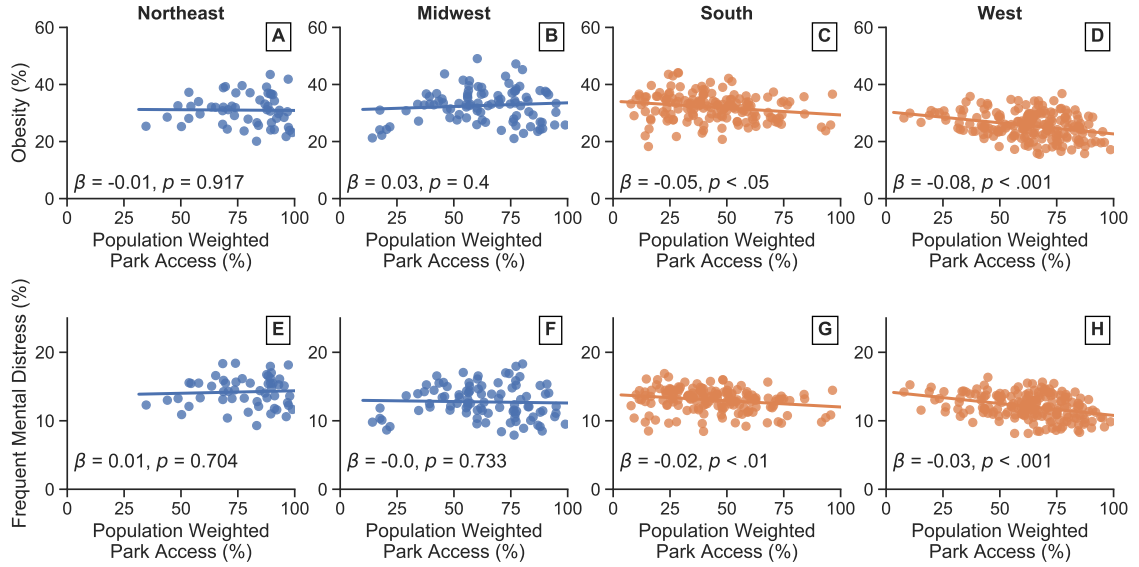


Figure 4.2: Regional scatterplots and regression lines for park access and health outcomes. Each point represents a city. Panels A-D: Park access and estimated obesity prevalence for each region. Panels E-H: Park access and estimated FMD prevalence for each region. Inset text displays the OLS regression coefficient  $\beta$  and p-value for population-weighted park access. Orange plots showed significant associations.

effects parameters, with the Midwest as reference (Table 4.3, Model 2). For FMD, the West had a significant negative fixed effect while the Northeast had a significant positive fixed effect, with the Midwest as reference (Table 4.3, Model 4).

#### 4.4.2 Tract Level Access & Equity

Having established the relationship between park access and health outcomes at the city level, we then looked at finer-scale neighborhood data to examine the equity of park access among social and economic groups. We found that relationships with park access varied widely among socioeconomic factors. Relationships were non-linear and not all in expected directions.

We found a u-shaped curve for the relationship between median income and park access. Average park access was highest at the low and high ends of the income range

Table 4.3: Regression models for Obesity & FMD.  
Regional coefficients are relative to reference region of the Midwest. \* $p < 0.05$

	Obesity (%)	Obesity (%)	FMD (%)	FMD (%)
	(1)	(2)	(3)	(4)
Park Access	-0.06* (-0.08,-0.03)	-0.04* (-0.06,-0.02)	-0.01* (-0.02,-0.005)	-0.02* (-0.03,-0.01)
Northeast		-1.02 (-2.71,0.66)		1.76* (1.11,2.40)
South		-1.36* (-2.70,-0.02)		-0.09 (-0.60,0.43)
West		-7.07* (-8.29,-5.85)		-0.69* (-1.15,-0.22)
Constant	32.77* (31.33,34.22)	35.14* (33.37,36.91)	13.49* (12.99,14.00)	13.97* (13.29,14.65)

(Fig. 4.3A). As percent population of White residents increased, we found a downward trend in park access (Fig. 4.3B). As percent population of black residents increased, park access tended to increase (Fig. 4.3C). As median age of a census tract increased, park access tended to decrease (Fig. 4.3D). The relationship between population with bachelor's degree and park access was u-shaped, similar to median income (Fig. 4.3E). The shapes of these curves varied regionally. Notably, the smoothed curves for the Northeast region were flatter than other regions, due to the more uniformly high park access across Northeastern census tracts.

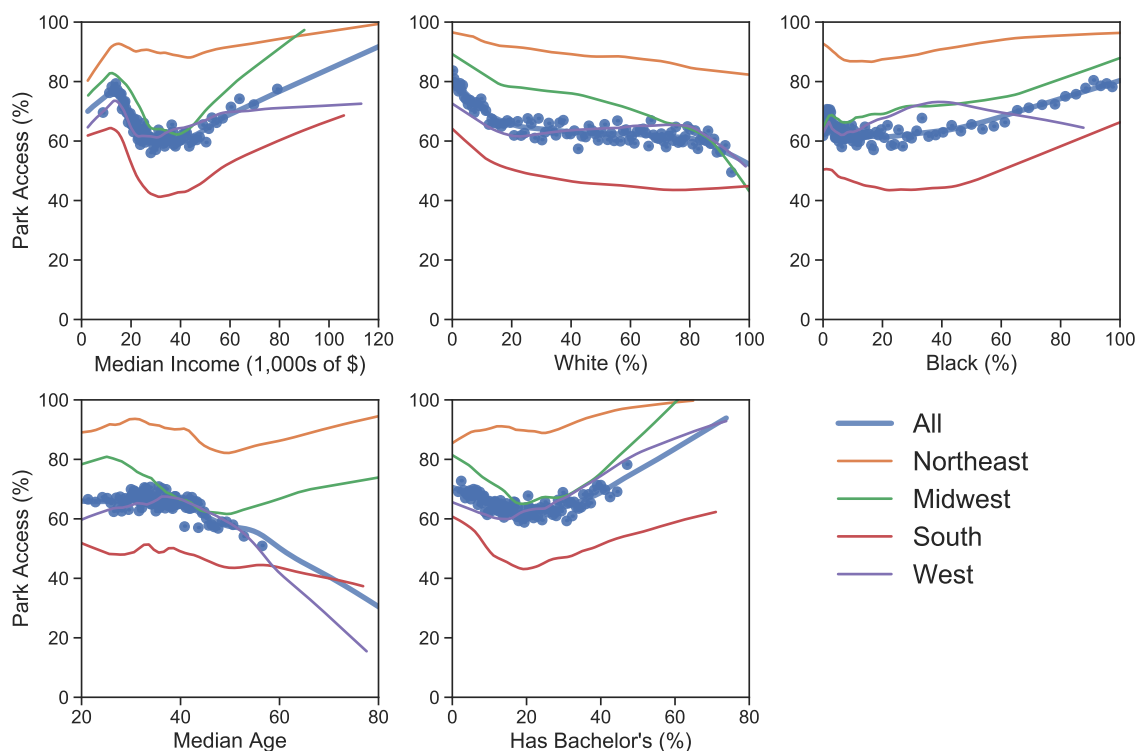


Figure 4.3: Relationships between park access and socioeconomic characteristics of neighborhoods. A: Median Income, B: Population White (%), C: Population Black (%), D: Median Age, E: Population with Bachelors (%). These data are from the US Census American Community Survey 5-year estimates. The plots represent smoothing across 26,501 census tracts for the full samples (blue curves). Blue dots show binned data along x-axis across all regions, with mean park access for that bin. Park access (%) is estimated based on area of census tract within 10-minutes walk of a park using ParkServe®.

## 4.5 Discussion

In this study, we synthesized several publicly available data sets to better understand the relationships among nature contact, human health, and socioeconomic equity in the US. Across the largest cities in the US, we found that as park access increases, there are lower rates of obesity and frequent mental distress (FMD). However, when we looked more closely at the major geographic regions of the US, only the South and West regions showed significant associations between park access and health outcomes. There are several possible explanations for this finding. Perhaps milder year-round climates result in greater opportunities for utilizing and activating the benefits of parks. On average, cities in the Northeast have the greatest population-weighted park access (78%). Therefore, it is possible that other factors are driving city to city differences in health outcomes in the Northeast, where park access is more ubiquitous. The significance of geographic region suggests that spatial context is related to these health outcomes; we suggest that future modeling efforts adapt more geographically informed approaches that take regional differences into account.

Our analysis of census tract level park access and socioeconomic and demographic variables suggested some non-intuitive patterns. Park access was greatest at the low and high ends of median income. This echoes some of the mixed results seen in prior studies of park access and income. Predominantly white census tracts tend to have lower park access on average. This may reflect different patterns of racial groups in the urban core and periphery of cities, where there may be lower park access within walking distance. Although there is evidence that low-income and minority populations have greater park accessibility, other studies have shown the parks in these neighborhoods are of lower quality (Rigolon et al., 2018). These inequities in park access and quality are related to socioeconomic factors and institutional capacity

(Leon-Moreta, Totaro, and Dixon, 2020). Future work could consider both park access and quality jointly to better assess the equity implications of urban parks for health outcomes.

Recent work has demonstrated that greenspace may have differential impacts on health throughout the life-course (Astell-Burt, Mitchell, and Hartig, 2014). We found that older populations have lower access to parks and may be missing out on potential health benefits of urban greenspace. Education status showed a non-linear relationship with park access. As population with a Bachelors increased from 0 to around 20%, park access decreased, and then began to increase (Fig. 4.2E). These non-linear relationships indicate that the link between park access and health is complex and nuanced.

Future efforts in this area of research would benefit from direct measures of population level health, rather than the synthetic estimates provided by the CDC 500 Cities data. While these types of data collection efforts are expensive, direct estimates of health would allow researchers to implement more robust modeling techniques that could compare the relative contribution of different population factors to health outcomes, and account for non-linear relationships between park access and socioeconomic and demographic factors. Currently available national health data limit our analysis because they rely on a modeling process that already takes income, race, education, and age into account. We need direct estimates of health outcomes in order to model the complex relationships between socioeconomic and environmental factors that impact health. In addition, we were only able to model park access and health at the city level due to the reliability of small-area estimates at the census tract level (Kong and Zhang, 2020).

This study, based on population-level cross-sectional data, has several additional limitations. We are unable to make causal claims about park access and health;



laboratory, field or randomized control trial experiments are better suited for causal inference. Studies at the single-city level, such as a recent analysis of vacant lot greening and mental health with random assignment, can provide insight into the causal pathways linking greenspace and health (South, Hohl, Kondo, MacDonald, and Branas, 2018). The CDC 500 Cities data are estimated based on survey data; this technique has limitations including known biases in self-reported health data (Nyholm et al., 2007; Stommel and Schoenborn, 2009). The Trust for Public Land may not capture all areas in a city that function as parks; greenspaces other than municipal parks (e.g., university campuses) are omitted. High resolution local studies using tree canopy data and local knowledge could allow for more nuanced investigations at the single city or neighborhood level. Localized studies can also take into account park attributes such as vegetation, biodiversity, amenities, and the local configuration of parks, all of which potentially moderate how a dose of nature impacts health (Roberts et al., 2019; Wang and Tassinary, 2019). In this study, we limited our focus to local parks within urban boundaries; some cities may have better access to larger open spaces such as national parks or wilderness areas.

Urban parks provide many benefits beyond their potential for improving health. Parks supply a suite of ecosystem services including air quality improvements, carbon storage, urban heat island reduction, protection from floods, and habitat for wildlife (Keeler et al., 2019; Kremer, Hamstead, and McPhearson, 2016). However, cities in the US have decreased their funding for parks over the last five decades (Rigolon et al., 2018). The monetary valuation of the health benefits of parks may help generate increased support for park funding. For example, researchers have found that nearby greenspaces can lead to reduced health care costs via increased physical activity and reduced use of prescription drugs (Helbich, Klein, Roberts, Hagedoorn, and Groenewegen, 2018; Sato, Inoue, Du, and Funk, 2019). Urban parks are much

more than a place to recreate; it is becoming clear that parks are also a critical piece of public health infrastructure.

## 4.6 References

- Astell-Burt, T., Mitchell, R., & Hartig, T. (2014). The association between green space and mental health varies across the lifecourse. a longitudinal study. *J Epidemiol Community Health*, 68(6), 578–583.  
DOI: [10.1136/jech-2013-203767](https://doi.org/10.1136/jech-2013-203767)
- Browning, M. H., & Rigolon, A. (2018). Do income, race and ethnicity, and sprawl influence the greenspace-human health link in city-level analyses? findings from 496 cities in the united states. *International journal of environmental research and public health*, 15(7), 1541.  
DOI: [10.3390/ijerph15071541](https://doi.org/10.3390/ijerph15071541)
- CDC. (2018a). 500 cities: Local data for better health. <https://www.cdc.gov/500cities>. Atlanta, Georgia: U.S. Department of Health, Human Services, Centers for Disease Control, and Prevention.
- CDC. (2018b). Behavioral risk factor surveillance system survey data. <https://www.cdc.gov/brfss/index.html>. Atlanta, Georgia: U.S. Department of Health, Human Services, Centers for Disease Control, and Prevention.
- Fan, Y., Das, K. V., & Chen, Q. (2011). Neighborhood green, social support, physical activity, and stress: Assessing the cumulative impact. *Health & place*, 17(6), 1202–1211.  
DOI: [10.1016/j.healthplace.2011.08.008](https://doi.org/10.1016/j.healthplace.2011.08.008)
- Frumkin, H., Bratman, G. N., Breslow, S. J., Cochran, B., Kahn Jr, P. H., Lawler, J. J., ... Wood, S. A. (2017). Nature contact and human health: A research agenda. *Environmental Health Perspectives*, 125(7), 1–18.  
DOI: [10.1289/EHP1663](https://doi.org/10.1289/EHP1663)
- Helbich, M., Klein, N., Roberts, H., Hagedoorn, P., & Groenewegen, P. P. (2018). More green space is related to less antidepressant prescription rates in the netherlands: A bayesian geoadditive quantile regression approach. *Environmental research*, 166, 290–297.  
DOI: [10.1016/j.envres.2018.06.01](https://doi.org/10.1016/j.envres.2018.06.01)

- Jacoby, W. G. (2000). Loess:: A nonparametric, graphical tool for depicting relationships between variables. *Electoral Studies*, 19(4), 577–613.  
DOI: [10.1016/S0261-3794\(99\)00028-1](https://doi.org/10.1016/S0261-3794(99)00028-1)
- Kabisch, N., Qureshi, S., & Haase, D. (2015). Human–environment interactions in urban green spaces—a systematic review of contemporary issues and prospects for future research. *Environmental Impact Assessment Review*, 50, 25–34.  
DOI: [10.1016/j.eiar.2014.08.007](https://doi.org/10.1016/j.eiar.2014.08.007)
- Keeler, B. L., Hamel, P., McPhearson, T., Hamann, M. H., Donahue, M. L., Prado, K. A. M., ... Finlay, J. C., et al. (2019). Social-ecological and technological factors moderate the value of urban nature. *Nature Sustainability*, 2(1), 29–38.  
DOI: [10.1038/s41893-018-0202-1](https://doi.org/10.1038/s41893-018-0202-1)
- Kong, A. Y., & Zhang, X. (2020). The use of small area estimates in place-based health research. *American Journal of Public Health*, (0), e1–e4.  
DOI: [10.2105/AJPH.2020.305611](https://doi.org/10.2105/AJPH.2020.305611)
- Kremer, P., Hamstead, Z. A., & McPhearson, T. (2016). The value of urban ecosystem services in new york city: A spatially explicit multicriteria analysis of landscape scale valuation scenarios. *Environmental Science & Policy*, 62, 57–68.  
DOI: [10.1016/j.envsci.2016.04.012](https://doi.org/10.1016/j.envsci.2016.04.012)
- Kuo, M. (2015). How might contact with nature promote human health? promising mechanisms and a possible central pathway. *Frontiers in psychology*, 6, 1093.  
DOI: [10.3389/fpsyg.2015.01093](https://doi.org/10.3389/fpsyg.2015.01093)
- Larson, L. R., Jennings, V., & Cloutier, S. A. (2016). Public parks and wellbeing in urban areas of the united states. *PLoS one*, 11(4).  
DOI: [10.1371/journal.pone.0153211](https://doi.org/10.1371/journal.pone.0153211)
- Leon-Moreta, A., Totaro, V. R., & Dixon, D. S. (2020). Social heterogeneity, local capacity, and urban parks: Evidence from us cities. *Cities*, 99, 102588.  
DOI: [10.1016/j.cities.2019.102588](https://doi.org/10.1016/j.cities.2019.102588)
- Markevych, I., Schoierer, J., Hartig, T., Chudnovsky, A., Hystad, P., Dzhambov, A. M., ... Fuertes, E. (2017). Exploring pathways linking greenspace to health: Theoretical and methodological guidance. *Environmental Research*, 158(June), 301–317.  
DOI: [10.1016/j.envres.2017.06.028](https://doi.org/10.1016/j.envres.2017.06.028)
- Moriarty, D. G., Zack, M. M., Holt, J. B., Chapman, D. P., & Safran, M. A. (2009). Geographic patterns of frequent mental distress: Us adults, 1993–2001 and 2003–2006. *American journal of preventive medicine*, 36(6), 497–505.  
DOI: [10.1016/j.amepre.2009.01.038](https://doi.org/10.1016/j.amepre.2009.01.038)

- Mullenbach, L. E., Mowen, A. J., & Baker, B. L. (2018). Peer reviewed: Assessing the relationship between a composite score of urban park quality and health. *Preventing chronic disease*, 15.  
DOI: [10.5888/pcd15.180033](https://doi.org/10.5888/pcd15.180033)
- Nesbitt, L., Meitner, M. J., Girling, C., Sheppard, S. R., & Lu, Y. (2019). Who has access to urban vegetation? a spatial analysis of distributional green equity in 10 us cities. *Landscape and Urban Planning*, 181, 51–79.  
DOI: [10.1016/j.landurbplan.2018.08.007](https://doi.org/10.1016/j.landurbplan.2018.08.007)
- Nyholm, M., Gullberg, B., Merlo, J., Lundqvist-Persson, C., Råstam, L., & Lindblad, U. (2007). The validity of obesity based on self-reported weight and height: Implications for population studies. *Obesity*, 15(1), 197–197.  
DOI: [10.1038/oby.2007.536](https://doi.org/10.1038/oby.2007.536)
- Rigolon, A., Browning, M., & Jennings, V. (2018). Inequities in the quality of urban park systems: An environmental justice investigation of cities in the united states. *Landscape and Urban Planning*, 178, 156–169.  
DOI: [10.1016/j.landurbplan.2018.05.026](https://doi.org/10.1016/j.landurbplan.2018.05.026)
- Roberts, H., Kellar, I., Conner, M., Gidlow, C., Kelly, B., Nieuwenhuijsen, M., & McEachan, R. (2019). Associations between park features, park satisfaction and park use in a multi-ethnic deprived urban area. *Urban Forestry & Urban Greening*, 46, 126485.  
DOI: [10.1016/j.ufug.2019.126485](https://doi.org/10.1016/j.ufug.2019.126485)
- Sato, M., Inoue, Y., Du, J., & Funk, D. C. (2019). Access to parks and recreational facilities, physical activity, and health care costs for older adults: Evidence from us counties. *Journal of Leisure Research*, 50(3), 220–238.  
DOI: [10.1080/00222216.2019.1583048](https://doi.org/10.1080/00222216.2019.1583048)
- Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and statistical modeling with python. In *9th python in science conference*.
- South, E. C., Hohl, B. C., Kondo, M. C., MacDonald, J. M., & Branas, C. C. (2018). Effect of greening vacant land on mental health of community-dwelling adults: A cluster randomized trial. *JAMA network open*, 1(3), e180298–e180298.  
DOI: [10.1001/jamanetworkopen.2018.0298](https://doi.org/10.1001/jamanetworkopen.2018.0298)
- Stommel, M., & Schoenborn, C. A. (2009). Accuracy and usefulness of bmi measures based on self-reported weight and height: Findings from the nhanes & nhis 2001-2006. *BMC public health*, 9(1), 421.  
DOI: [10.1186/1471-2458-9-421](https://doi.org/10.1186/1471-2458-9-421)
- The Trust for Public Land. (2019). Park serve. data retrieved from TPL, <https://www.tpl.org/parkserve/downloads>.

- Tsai, W.-L., Davis, A. J., & Jackson, L. E. (2019). Associations between types of greenery along neighborhood roads and weight status in different climates. *Urban forestry & urban greening*, 41, 104–117.  
DOI: [10.1016/j.ufug.2019.03.011](https://doi.org/10.1016/j.ufug.2019.03.011)
- Tsai, W.-L., McHale, M. R., Jennings, V., Marquet, O., Hipp, J. A., Leung, Y.-F., & Floyd, M. F. (2018). Relationships between characteristics of urban green land cover and mental health in us metropolitan areas. *International journal of environmental research and public health*, 15(2), 340.  
DOI: [10.3390/ijerph15020340](https://doi.org/10.3390/ijerph15020340)
- U.S. Census Bureau. (2017). American Community Survey (ACS): 5-Year Estimates 2013-2017. <https://www.census.gov/programs-surveys/acs>.
- Van den Berg, M., Wendel-Vos, W., van Poppel, M., Kemper, H., van Mechelen, W., & Maas, J. (2015). Health benefits of green spaces in the living environment: A systematic review of epidemiological studies. *Urban Forestry & Urban Greening*, 14(4), 806–816.  
DOI: [10.1016/j.ufug.2015.07.008](https://doi.org/10.1016/j.ufug.2015.07.008)
- Wang, H., & Tassinary, L. G. (2019). Effects of greenspace morphology on mortality at the neighbourhood level: A cross-sectional ecological study. *The Lancet Planetary Health*, 3(11), e460–e468.  
DOI: [10.1016/S2542-5196\(19\)30217-7](https://doi.org/10.1016/S2542-5196(19)30217-7)
- Wen, M., Zhang, X., Harris, C. D., Holt, J. B., & Croft, J. B. (2013). Spatial disparities in the distribution of parks and green spaces in the usa. *Annals of Behavioral Medicine*, 45(suppl.1), S18–S27.  
DOI: [10.1007/s12160-012-9426-x](https://doi.org/10.1007/s12160-012-9426-x)
- West, S. T., Shores, K. A., & Mudd, L. M. (2012). Association of available parkland, physical activity, and overweight in america's largest cities. *Journal of Public Health Management and Practice*, 18(5), 423–430.  
DOI: [10.1097/PHH.0b013e318238ea27](https://doi.org/10.1097/PHH.0b013e318238ea27)
- Zhang, L., Tan, P. Y., & Diehl, J. A. (2017). A conceptual framework for studying urban green spaces effects on health. *Journal of Urban Ecology*, 3(1), jux015.  
DOI: [10.1093/jue/jux015](https://doi.org/10.1093/jue/jux015)

# Chapter 5: Conclusion

## 5.1 Synthesis & Next Steps

Researchers across several disciplines have made significant progress in understanding the relationship between nature contact and human health. While the results from these studies have received some media attention, using greenspace for health promotion is still lagging in most places. The results of this dissertation help fill important gaps in our understanding of the health benefits of urban parks at the city level. In addition, the methods developed in this dissertation can produce results that can clearly communicate the health benefits of parks to decision makers and citizens.

In Chapters 2 and 3, I collected tweets within urban parks and analyzed the words people wrote to quantify the mental benefit of park visitation. In Chapter 2, I found that tweets in San Francisco parks were happier than tweets in the hours leading up to those park visits. Tweets remained happier for up to four hours following a park visit. The increase in tweet happiness during park visits was equivalent with the increase across all tweets on Christmas day, the happiest day of the year on Twitter. Contextualizing the benefits of nature contact in this way can facilitate clearer communication of these results to broad audiences. In Chapter 3, I expanded this analysis to the largest 25 cities in the US. I found that tweets in parks were happier compared to tweets outside of parks in all 25 cities. However, park investment and a composite score of park quality did not explain inter-city differences in the benefits of park visits. In further studies, I plan to investigate what factors are driving the differences in inter-city park sentiment.

Tweets in the largest parks exhibited the greatest increase in happiness compared

to out-of-park tweets in both Chapters 2 and 3. Larger parks may be particularly well-suited for providing refuge from the stressors of the urban landscape. These larger parks may also contain more landscape types and greater biodiversity, both of which may activate restorative processes related to nature contact. People may also choose to spend more time in larger parks, increasing the potential for improving mood and reducing stress. With more detailed data at the park-level, researchers could test which other park features are the most important for providing health benefits.

The methods developed in these chapters could be applied to other geographic and cultural contexts to better understand the contribution of nature contact to human health. While finding sufficient tweet density outside of cities may be challenging, it would be revealing to quantify the benefits of nature contact for people in exurban, suburban, and rural areas. In the nature contact and health literature, countries outside of North America and Western Europe have received less attention historically. Expanding the analysis to additional languages and countries would provide insight into how different cultures benefit from nature contact.

In Chapter 4, I expanded the geographic scope of analysis to 500 cities across the US and synthesized several publicly available data sets to investigate the association between park access and health. In these results, park access is positively associated with lower rates of obesity and frequent mental distress, though this association was not significant in all regions of the country. Park access at the neighborhood level (as estimated within census tracts) exhibited non-linear associations with income. This result is non-intuitive because the assumption has been that poorer neighborhoods typically have worse park access. However, measuring park access without accounting for park quality may be masking the true distribution of opportunities for nature contact that actually promote health. In future work, I hope to find ways to combine

park quantity, access, and quality to gain a more nuanced understanding of the relationship between nature contact, health, and equity.

## 5.2 Broader Impacts

The findings of this research are directed at decision-makers at the city level. While the overall utility of parks is not usually contested, park systems have recently not been a funding priority. Framing parks as public health infrastructure may help garner political support and funding to maintain and improve existing parks or even construct new ones in communities with low access or park quality. Municipal parks departments could consider using data from social media such as Twitter to better understand what activities and amenities are most important in the parks they manage. These data sets are readily available and would not require conducting surveys, which eliminates a time and resource barrier for decision makers. The results of Chapters 3 demonstrated significant variation in park investment per capita at the city level and access to parks at the neighborhood or community level. Tools such as ParkServe® can support parks departments in finding the places with the greatest need for parks. Pairing spatial support tools like ParkServe® with estimates of well-being from social media is a potentially powerful combination for better understanding local needs.

In health care, there have been some efforts, such as park prescriptions, to recommend park visits for health promotion. These programs are promising because they do not require significant investment—people can use existing park infrastructure. Public health officials and healthcare practitioners can add park prescriptions to integrative treatment approaches for both obesity and mental health. In this dissertation, visits to larger parks were shown to most strongly promote mental health



benefits, compared with smaller parks. However, all park types conferred mental health benefits, and healthcare practitioners can tailor their park prescriptions to individual preferences and park access.

The way people use and access urban parks and other public spaces has been brought to the forefront during the COVID-19 pandemic. Our largest urban centers such as New York City have been the epicenters of the pandemic. Society-wide shutdowns and social distancing to ‘flatten the curve’ have included the closure of public spaces. At the same time, the pandemic is having a massive impact on mental health across society. While spending time in nature is a promising avenue to ease this burden, it must be balanced with concerns about spreading the virus. As state governments begin to relax stay-at-home orders and reopen the economy, we are going to see new norms around the use of urban parks. We must find ways for urban residents to safely use parks in these novel circumstances— it is critical for supporting both collective and individual health. Perhaps this reminder of the importance of parks can galvanize future efforts into reconnecting people with nature in cities, so that more people can access and benefit from the many health benefits of nature contact.

## Chapter 6: Bibliography

- Ambrey, C. L. (2016). An investigation into the synergistic wellbeing benefits of greenspace and physical activity: Moving beyond the mean. *Urban Forestry & Urban Greening*, 19, 7–12.  
DOI: [10.1016/j.ufug.2016.06.020](https://doi.org/10.1016/j.ufug.2016.06.020)
- Astell-Burt, T., & Feng, X. (2019). Association of urban green space with mental health and general health among adults in Australia. *JAMA network open*, 2(7), e198209–e198209.  
DOI: [10.1001/jamanetworkopen.2019.8209](https://doi.org/10.1001/jamanetworkopen.2019.8209)
- Astell-Burt, T., Mitchell, R., & Hartig, T. (2014). The association between green space and mental health varies across the lifecourse. a longitudinal study. *J Epidemiol Community Health*, 68(6), 578–583.  
DOI: [10.1136/jech-2013-203767](https://doi.org/10.1136/jech-2013-203767)
- Bakolis, I., Hammoud, R., Smythe, M., Gibbons, J., Davidson, N., Tognin, S., & Mechelli, A. (2018). Urban mind: Using smartphone technologies to investigate the impact of nature on mental well-being in real time. *BioScience*, 68(2), 134–145.  
DOI: [10.1093/biosci/bix149](https://doi.org/10.1093/biosci/bix149)
- Baylis, P., Obradovich, N., Kryvasheyeu, Y., Chen, H., Coviello, L., Moro, E., . . . Fowler, J. H. (2018). Weather impacts expressed sentiment. *PloS one*, 13(4).  
DOI: [10.1371/journal.pone.0195750](https://doi.org/10.1371/journal.pone.0195750)
- Bell, S. L., Phoenix, C., Lovell, R., & Wheeler, B. W. (2014). Green space, health and wellbeing: making space for individual agency. *Health & Place*, 30, 287–292.  
DOI: [10.1016/j.healthplace.2014.10.005](https://doi.org/10.1016/j.healthplace.2014.10.005)
- Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., & Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3-4), 239–258.  
DOI: [10.1016/j.isprsjprs.2003.10.002](https://doi.org/10.1016/j.isprsjprs.2003.10.002)
- Berto, R. (2014). The Role of Nature in Coping with Psycho-Physiological Stress: A Literature Review on Restorativeness. *Behavioral Sciences*, 4(4), 394–409.  
DOI: [10.3390/bs4040394](https://doi.org/10.3390/bs4040394)
- Bettencourt, L. M., Lobo, J., Helbing, D., Kühnert, C., & West, G. B. (2007). Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the national academy of sciences*, 104(17), 7301–7306.  
DOI: [10.1073/pnas.0610172104](https://doi.org/10.1073/pnas.0610172104)

- Botzat, A., Fischer, L. K., & Kowarik, I. (2016). Unexploited opportunities in understanding liveable and biodiverse cities. A review on urban biodiversity perception and valuation. *Global Environmental Change*, 39, 220–233.  
DOI: [10.1016/j.gloenvcha.2016.04.008](https://doi.org/10.1016/j.gloenvcha.2016.04.008)
- Boyd, F., White, M. P., Bell, S. L., & Burt, J. (2018). Who doesn't visit natural environments for recreation and why: A population representative analysis of spatial, individual and temporal factors among adults in england. *Landscape and urban planning*, 175, 102–113.  
DOI: [10.1016/j.landurbplan.2018.03.016](https://doi.org/10.1016/j.landurbplan.2018.03.016)
- Bratman, G., Anderson, C. B., Berman, M. G., Cochran, B., De Vries, S., Flanders, J., ... Hartig, T., et al. (2019). Nature and mental health: An ecosystem service perspective. *Science advances*, 5(7), eaax0903.
- Bratman, G., Daily, G. C., Levy, B. J., & Gross, J. J. (2015). The benefits of nature experience: Improved affect and cognition. *Landscape and Urban Planning*, 138, 41–50.  
DOI: [10.1016/j.landurbplan.2015.02.005](https://doi.org/10.1016/j.landurbplan.2015.02.005)
- Bratman, G., Hamilton, J. P., Hahn, K. S., Daily, G. C., & Gross, J. J. (2015). Nature experience reduces rumination and subgenual prefrontal cortex activation. *Proceedings of the National Academy of Sciences*, 112(28), 8567–8572.  
DOI: [10.1073/pnas.1510459112](https://doi.org/10.1073/pnas.1510459112)
- Browning, M. H., & Rigolon, A. (2018). Do income, race and ethnicity, and sprawl influence the greenspace-human health link in city-level analyses? findings from 496 cities in the united states. *International journal of environmental research and public health*, 15(7), 1541.  
DOI: [10.3390/ijerph15071541](https://doi.org/10.3390/ijerph15071541)
- Buckley, R. C., & Brough, P. (2017). Economic value of parks via human mental health: An analytical framework. *Frontiers in Ecology and Evolution*, 5, 16.  
DOI: [doi.org/10.3389/fevo.2017.00016](https://doi.org/10.3389/fevo.2017.00016)
- Carrus, G., Scopelliti, M., Laforteza, R., Colangelo, G., Ferrini, F., Salbitano, F., ... Sanesi, G. (2015). Go greener, feel better? The positive effects of biodiversity on the well-being of individuals visiting urban and peri-urban green areas. *Landscape and Urban Planning*, 134, 221–228.  
DOI: [10.1016/j.landurbplan.2014.10.022](https://doi.org/10.1016/j.landurbplan.2014.10.022)
- CDC. (2018a). 500 cities: Local data for better health. <https://www.cdc.gov/500cities>. Atlanta, Georgia: U.S. Department of Health, Human Services, Centers for Disease Control, and Prevention.

- CDC. (2018b). Behavioral risk factor surveillance system survey data. <https://www.cdc.gov/brfss/index.html>. Atlanta, Georgia: U.S. Department of Health, Human Services, Centers for Disease Control, and Prevention.
- Clark, N. E., Lovell, R., Wheeler, B. W., Higgins, S. L., Depledge, M. H., & Norris, K. (2014). Biodiversity, cultural pathways, and human health: A framework. *Trends in ecology & evolution*, 29(4), 198–204.  
DOI: [10.1016/j.tree.2014.01.009](https://doi.org/10.1016/j.tree.2014.01.009)
- Cody, E. M., Reagan, A. J., Mitchell, L., Dodds, P. S., & Danforth, C. M. (2015). Climate change sentiment on Twitter: An unsolicited public opinion poll. *PLOS ONE*, 10(8), e0136092.  
DOI: [10.1371/journal.pone.0136092](https://doi.org/10.1371/journal.pone.0136092)
- Crnic, M., & Kondo, M. C. (2019). Nature rx: Reemergence of pediatric nature-based therapeutic programs from the late 19th and early 20th centuries. *American Public Health Association*, 109(10).  
DOI: [10.2105/AJPH.2019.305204](https://doi.org/10.2105/AJPH.2019.305204)
- Curiel, R. P., Cresci, S., Muntean, C. I., & Bishop, S. R. (2020). Crime and its fear in social media. *Palgrave Communications*, 6(1), 1–12.  
DOI: [10.1057/s41599-020-0430-7](https://doi.org/10.1057/s41599-020-0430-7)
- Dodds, P. S., & Danforth, C. M. (2010). Measuring the happiness of large-scale written expression: Songs, blogs, and presidents. *Journal of Happiness Studies*, 11(4), 441–456.  
DOI: [10.1007/s10902-009-9150-9](https://doi.org/10.1007/s10902-009-9150-9)
- Dodds, P. S., Harris, K. D., Kloumann, I. M., Bliss, C. A., & Danforth, C. M. (2011). Temporal patterns of happiness and information in a global social network: Hedonometrics and twitter. *PloS one*, 6(12).  
DOI: [10.1371/journal.pone.0026752](https://doi.org/10.1371/journal.pone.0026752)
- Elmqvist, T., Setälä, H., Handel, S. N., van der Ploeg, S., Aronson, J., Blignaut, J. N., ... de Groot, R. (2015). Benefits of restoring ecosystem services in urban areas. *Current Opinion in Environmental Sustainability*, 14, 101–108.  
DOI: [10.1016/j.cosust.2015.05.001](https://doi.org/10.1016/j.cosust.2015.05.001)
- Engemann, K., Pedersen, C. B., Arge, L., Tsirogiannis, C., Mortensen, P. B., & Svenning, J.-C. (2019). Residential green space in childhood is associated with lower risk of psychiatric disorders from adolescence into adulthood. *Proceedings of the national academy of sciences*, 116(11), 5188–5193.  
DOI: [10.1073/pnas.1807504116](https://doi.org/10.1073/pnas.1807504116)
- Fan, Y., Das, K. V., & Chen, Q. (2011). Neighborhood green, social support, physical activity, and stress: Assessing the cumulative impact. *Health & place*, 17(6),

1202–1211.

DOI: [10.1016/j.healthplace.2011.08.008](https://doi.org/10.1016/j.healthplace.2011.08.008)

Fischer, L. K., Honold, J., Cvejić, R., Delshammar, T., Hilbert, S., Laforтеzza, R., ... Kowarik, I. (2018). Beyond green: Broad support for biodiversity in multicultural European cities. *Global Environmental Change*, 49(February), 35–45.

DOI: [10.1016/j.gloenvcha.2018.02.001](https://doi.org/10.1016/j.gloenvcha.2018.02.001)

Fong, K. C., Hart, J. E., & James, P. (2018). A review of epidemiologic studies on greenness and health: updated literature through 2017. *Current Environmental Health Reports*, 1–11.

DOI: [10.1007/s40572-018-0179-y](https://doi.org/10.1007/s40572-018-0179-y)

Frumkin, H., Bratman, G. N., Breslow, S. J., Cochran, B., Kahn Jr, P. H., Lawler, J. J., ... Wood, S. A. (2017). Nature contact and human health: A research agenda. *Environmental Health Perspectives*, 125(7), 1–18.

DOI: [10.1289/EHP1663](https://doi.org/10.1289/EHP1663)

Gascon, M., Triguero-Mas, M., Martínez, D., Dadvand, P., Forns, J., Plasència, A., & Nieuwenhuijsen, M. (2015). Mental health benefits of long-term exposure to residential green and blue spaces: a systematic review. *International Journal of Environmental Research and Public Health*, 12(4), 4354–4379.

DOI: [10.3390/ijerph120404354](https://doi.org/10.3390/ijerph120404354)

Gascon, M., Triguero-Mas, M., Martínez, D., Dadvand, P., Rojas-Rueda, D., Plasència, A., & Nieuwenhuijsen, M. J. (2016). Residential green spaces and mortality: A systematic review. *Environment International*, 86, 60–67.

DOI: [10.1016/j.envint.2015.10.013](https://doi.org/10.1016/j.envint.2015.10.013)

Gidlow, C. J., Jones, M. V., Hurst, G., Masterson, D., Clark-Carter, D., Tarvainen, M. P., ... Nieuwenhuijsen, M. (2016). Where to put your best foot forward: Psycho-physiological responses to walking in natural and urban environments. *Journal of Environmental Psychology*, 45, 22–29.

DOI: [10.1016/j.jenvp.2015.11.003](https://doi.org/10.1016/j.jenvp.2015.11.003)

Gómez-Baggethun, E., & Barton, D. N. (2013). Classifying and valuing ecosystem services for urban planning. *Ecological Economics*, 86, 235–245.

DOI: [10.1016/j.ecolecon.2012.08.019](https://doi.org/10.1016/j.ecolecon.2012.08.019)

Gourevitch, M. N., Athens, J. K., Levine, S. E., Kleiman, N., & Thorpe, L. E. (2019). City-level measures of health, health determinants, and equity to foster population health improvement: The city health dashboard. *American journal of public health*, 109(4), 585–592.

DOI: [10.2105/AJPH.2018.304903](https://doi.org/10.2105/AJPH.2018.304903)

- Greenwood, S., Perrin, A., & Duggan, M. (2016). Demographics of social media users in 2016. <https://www.pewresearch.org/>.
- Harnik, P., Martin, A., & Barnhart, K. (2015). 2015 City Park Facts. <https://www.tpl.org>. Trust for Public land.
- Hartig, T., Mitchell, R., de Vries, S., & Frumkin, H. (2014). Nature and Health. *Annual Review of Public Health*, 35(1), 207–228.  
DOI: [10.1146/annurev-publhealth-032013-182443](https://doi.org/10.1146/annurev-publhealth-032013-182443)
- Hedonometer.org. (2020). Hedonometer, from The Vermont Complex Systems Center & Computational Story Lab. <http://hedonometer.org/>.
- Helbich, M., Klein, N., Roberts, H., Hagedoorn, P., & Groenewegen, P. P. (2018). More green space is related to less antidepressant prescription rates in the netherlands: A bayesian geoadditive quantile regression approach. *Environmental research*, 166, 290–297.  
DOI: [10.1016/j.envres.2018.06.01](https://doi.org/10.1016/j.envres.2018.06.01)
- Hofmann, M., Young, C., Binz, T., Baumgartner, M., & Bauer, N. (2017). Contact to nature benefits health: Mixed effectiveness of different mechanisms. *International Journal of Environmental Research and Public Health*, 15(1), 31.  
DOI: [10.3390/ijerph15010031](https://doi.org/10.3390/ijerph15010031)
- Houlden, V., Weich, S., de Albuquerque, J. P., Jarvis, S., & Rees, K. (2018). The relationship between greenspace and the mental wellbeing of adults: A systematic review. *PloS one*, 13(9).  
DOI: [10.1371/journal.pone.0203000](https://doi.org/10.1371/journal.pone.0203000)
- Hoyle, H., Hitchmough, J., & Jorgensen, A. (2017). All about the ‘wow factor’? The relationships between aesthetics, restorative effect and perceived biodiversity in designed urban planting. *Landscape and Urban Planning*, 164(2017), 109–123.  
DOI: [10.1016/j.landurbplan.2017.03.011](https://doi.org/10.1016/j.landurbplan.2017.03.011)
- Hoyle, H., Jorgensen, A., & Hitchmough, J. D. (2019). What determines how we see nature? Perceptions of naturalness in designed urban green spaces. *People and Nature*, (April).  
DOI: [10.1002/pan3.19](https://doi.org/10.1002/pan3.19)
- Jacoby, W. G. (2000). Loess:: A nonparametric, graphical tool for depicting relationships between variables. *Electoral Studies*, 19(4), 577–613.  
DOI: [10.1016/S0261-3794\(99\)00028-1](https://doi.org/10.1016/S0261-3794(99)00028-1)
- James, P., Banay, R. F., Hart, J. E., & Laden, F. (2015). A review of the health benefits of greenness. *Current epidemiology reports*, 2(2), 131–142.  
DOI: [10.1007/s40471-015-0044-6](https://doi.org/10.1007/s40471-015-0044-6)

- Kabisch, N., Qureshi, S., & Haase, D. (2015). Human–environment interactions in urban green spaces—a systematic review of contemporary issues and prospects for future research. *Environmental Impact Assessment Review*, 50, 25–34.  
DOI: [10.1016/j.eiar.2014.08.007](https://doi.org/10.1016/j.eiar.2014.08.007)
- Kaplan, R., & Kaplan, S. (1989). *The experience of nature: A psychological perspective*. CUP Archive.
- Kaplan, S. (1995). The restorative benefits of nature: Toward an integrative framework. *Journal of Environmental Psychology*, 15(3), 169–182.  
DOI: [10.1016/0272-4944\(95\)90001-2](https://doi.org/10.1016/0272-4944(95)90001-2)
- Keeler, B. L., Hamel, P., McPhearson, T., Hamann, M. H., Donahue, M. L., Prado, K. A. M., ... Finlay, J. C., et al. (2019). Social-ecological and technological factors moderate the value of urban nature. *Nature Sustainability*, 2(1), 29–38.  
DOI: [10.1038/s41893-018-0202-1](https://doi.org/10.1038/s41893-018-0202-1)
- Kellert, S. R., & Wilson, E. O. (1995). *The biophilia hypothesis*. Island Press.
- Kimpton, A., Corcoran, J., & Wickes, R. (2017). Greenspace and crime: An analysis of greenspace types, neighboring composition, and the temporal dimensions of crime. *Journal of Research in Crime and Delinquency*, 54(3), 303–337.  
DOI: [10.1177/0022427816666309](https://doi.org/10.1177/0022427816666309)
- Kong, A. Y., & Zhang, X. (2020). The use of small area estimates in place-based health research. *American Journal of Public Health*, (0), e1–e4.  
DOI: [10.2105/AJPH.2020.305611](https://doi.org/10.2105/AJPH.2020.305611)
- Krabbandam, L., van Vugt, M., Conus, P., Söderström, O., Empson, L. A., van Os, J., & Fett, A.-K. J. (2020). Understanding urbanicity: How interdisciplinary methods help to unravel the effects of the city on mental health. *Psychological Medicine*, 1–12.  
DOI: [10.1017/S0033291720000355](https://doi.org/10.1017/S0033291720000355)
- Kremer, P., Hamstead, Z. A., & McPhearson, T. (2016). The value of urban ecosystem services in new york city: A spatially explicit multicriteria analysis of landscape scale valuation scenarios. *Environmental Science & Policy*, 62, 57–68.  
DOI: [10.1016/j.envsci.2016.04.012](https://doi.org/10.1016/j.envsci.2016.04.012)
- Kuo, M. (2015). How might contact with nature promote human health? promising mechanisms and a possible central pathway. *Frontiers in psychology*, 6, 1093.  
DOI: [10.3389/fpsyg.2015.01093](https://doi.org/10.3389/fpsyg.2015.01093)
- Larson, L. R., Jennings, V., & Cloutier, S. A. (2016). Public parks and wellbeing in urban areas of the united states. *PLoS one*, 11(4).  
DOI: [10.1371/journal.pone.0153211](https://doi.org/10.1371/journal.pone.0153211)

- Leon-Moreta, A., Totaro, V. R., & Dixon, D. S. (2020). Social heterogeneity, local capacity, and urban parks: Evidence from us cities. *Cities*, 99, 102588.  
DOI: [10.1016/j.cities.2019.102588](https://doi.org/10.1016/j.cities.2019.102588)
- Li, W., Kelsey, J. L., Zhang, Z., Lemon, S. C., Mezgebu, S., Boddie-Willis, C., & Reed, G. W. (2009). Small-area estimation and prioritizing communities for obesity control in massachusetts. *American journal of public health*, 99(3), 511–519.  
DOI: [10.2105/ajph.2008.137364](https://doi.org/10.2105/ajph.2008.137364)
- Lim, K. H., Lee, K. E., Kendal, D., Rashidi, L., Naghizade, E., Winter, S., & Vasardani, M. (2018). The grass is greener on the other side: Understanding the effects of green spaces on Twitter user sentiments. In *Companion of the the web conference 2018 on the web conference 2018 - www '18* (pp. 275–282).  
DOI: [10.1145/3184558.3186337](https://doi.org/10.1145/3184558.3186337)
- Louv, R. (2011). *The nature principle: Human restoration and the end of nature-deficit disorder*. Algonquin Books.
- Maas, J., Verheij, R. A., Groenewegen, P. P., de Vries, S., & Spreeuwenberg, P. (2006). Green space, urbanity, and health: how strong is the relation? *Journal of Epidemiology and Community Health*, 60(7), 587–592.  
DOI: [10.1136/jech.2005.043125](https://doi.org/10.1136/jech.2005.043125)
- MacKerron, G., & Mourato, S. (2013). Happiness is greater in natural environments. *Global Environmental Change*, 23(5), 992–1000.  
DOI: [10.1016/j.gloenvcha.2013.03.010](https://doi.org/10.1016/j.gloenvcha.2013.03.010)
- Markevych, I., Schoierer, J., Hartig, T., Chudnovsky, A., Hystad, P., Dzhambov, A. M., ... Fuertes, E. (2017). Exploring pathways linking greenspace to health: Theoretical and methodological guidance. *Environmental Research*, 158(June), 301–317.  
DOI: [10.1016/j.envres.2017.06.028](https://doi.org/10.1016/j.envres.2017.06.028)
- Martens, D., Gutscher, H., & Bauer, N. (2011). Walking in “wild” and “tended” urban forests: The impact on psychological well-being. *Journal of Environmental Psychology*, 31(1), 36–44.  
DOI: [10.1016/j.jenvp.2010.11.001](https://doi.org/10.1016/j.jenvp.2010.11.001)
- Mavoa, S., Davern, M., Breed, M., & Hahs, A. (2019). Higher levels of greenness and biodiversity associate with greater subjective wellbeing in adults living in melbourne, australia. *Health & place*, 57, 321–329.  
DOI: [10.1016/j.healthplace.2019.05.006](https://doi.org/10.1016/j.healthplace.2019.05.006)
- McDonald, R. I., Beatley, T., & Elmqvist, T. (2018). The green soul of the concrete jungle: the urban century, the urban psychological penalty, and the role of nature. *Sustainable Earth*, 1x'1(1), 3.  
DOI: [10.1186/s42055-018-0002-5](https://doi.org/10.1186/s42055-018-0002-5)



- McEwan, K., Richardson, M., Brindley, P., Sheffield, D., Tait, C., Johnson, S., ... Ferguson, F. J. (2019). Shmapped: development of an app to record and promote the well-being benefits of noticing urban nature. *Translational Behavioral Medicine*, 1–11.  
DOI: [10.1093/tbm/ibz027](https://doi.org/10.1093/tbm/ibz027)
- McMahan, E. A., & Estes, D. (2015). The effect of contact with natural environments on positive and negative affect: A meta-analysis. *The Journal of Positive Psychology*, 10(6), 507–519.  
DOI: [10.1080/17439760.2014.994224](https://doi.org/10.1080/17439760.2014.994224)
- Mitchell, L., Frank, M. R., Harris, K. D., Dodds, P. S., & Danforth, C. M. (2013). The geography of happiness: Connecting twitter sentiment and expression, demographics, and objective characteristics of place. *PLoS ONE*, 8(5), e64417.  
DOI: [10.1371/journal.pone.0064417](https://doi.org/10.1371/journal.pone.0064417)
- Moore, F. C., Obradovich, N., Lehner, F., & Baylis, P. (2019). Rapidly declining remarkability of temperature anomalies may obscure public perception of climate change. *Proceedings of the National Academy of Sciences*, 116(11), 4905–4910.  
DOI: [10.1073/pnas.1816541116](https://doi.org/10.1073/pnas.1816541116)
- Moriarty, D. G., Zack, M. M., Holt, J. B., Chapman, D. P., & Safran, M. A. (2009). Geographic patterns of frequent mental distress: Us adults, 1993–2001 and 2003–2006. *American journal of preventive medicine*, 36(6), 497–505.  
DOI: [10.1016/j.amepre.2009.01.038](https://doi.org/10.1016/j.amepre.2009.01.038)
- Mueller, J. T., Park, S. Y., & Mowen, A. J. (2019). The relationship between self-rated health and local government spending on parks and recreation in the united states from 1997 to 2012. *Preventive medicine reports*, 13, 105–112.  
DOI: [10.1016/j.pmedr.2018.11.018](https://doi.org/10.1016/j.pmedr.2018.11.018)
- Mullenbach, L. E., Mowen, A. J., & Baker, B. L. (2018). Peer reviewed: Assessing the relationship between a composite score of urban park quality and health. *Preventing chronic disease*, 15.  
DOI: [10.5888/pcd15.180033](https://doi.org/10.5888/pcd15.180033)
- Murray, C. J., Vos, T., Lozano, R., Naghavi, M., Flaxman, A. D., Michaud, C., ... Abdalla, S., et al. (2012). Disability-adjusted life years (dalys) for 291 diseases and injuries in 21 regions, 1990–2010: A systematic analysis for the global burden of disease study 2010. *The lancet*, 380(9859), 2197–2223.  
DOI: [10.1016/S0140-6736\(12\)61689-4](https://doi.org/10.1016/S0140-6736(12)61689-4)
- Nesbitt, L., Meitner, M. J., Girling, C., Sheppard, S. R., & Lu, Y. (2019). Who has access to urban vegetation? a spatial analysis of distributional green equity in

- 10 us cities. *Landscape and Urban Planning*, 181, 51–79.  
DOI: [10.1016/j.landurbplan.2018.08.007](https://doi.org/10.1016/j.landurbplan.2018.08.007)
- Nielsen. (2017). The Nielsen Total Audience Report — Q2 2017. <https://www.nielsen.com>.
- Nieuwenhuijsen, M. J., Khreis, H., Triguero-Mas, M., Gascon, M., & Dadvand, P. (2017). Fifty shades of green. *Epidemiology*, 28(1), 63–71.  
DOI: [10.1097/EDE.0000000000000549](https://doi.org/10.1097/EDE.0000000000000549)
- Nyholm, M., Gullberg, B., Merlo, J., Lundqvist-Persson, C., Råstam, L., & Lindblad, U. (2007). The validity of obesity based on self-reported weight and height: Implications for population studies. *Obesity*, 15(1), 197–197.  
DOI: [10.1038/oby.2007.536](https://doi.org/10.1038/oby.2007.536)
- O’Neil-Dunne, J., Pelletier, K., MacFaden, S., Troy, A., & Grove, J. (2009). Object-based high-resolution land-cover mapping. In *2009 17th international conference on geoinformatics* (pp. 1–6).  
DOI: [10.1109/GEOINFORMATICS.2009.5293435](https://doi.org/10.1109/GEOINFORMATICS.2009.5293435)
- Ohly, H., White, M. P., Wheeler, B. W., Bethel, A., Ukoumunne, O. C., Nikolaou, V., & Garside, R. (2016). Attention Restoration Theory: A systematic review of the attention restoration potential of exposure to natural environments. *Journal of Toxicology and Environmental Health, Part B*, 19(7), 305–343.  
DOI: [10.1080/10937404.2016.1196155](https://doi.org/10.1080/10937404.2016.1196155)
- Park and Open Space Map. (2016). <https://data.sfgov.org/>. San Francisco Recreation & Parks.
- Pennebaker, J. W. (2011). *The Secret Life of Pronouns: What our Words Say about us*.  
DOI: [10.1093/llc/fqt006](https://doi.org/10.1093/llc/fqt006)
- Plunz, R. A., Zhou, Y., Carrasco Vintimilla, M. I., Mckeown, K., Yu, T., Uguccioni, L., & Sutto, M. P. (2019). Twitter sentiment in New York City parks as measure of well-being. *Landscape and Urban Planning*, 189(May), 235–246.  
DOI: [10.1016/j.landurbplan.2019.04.024](https://doi.org/10.1016/j.landurbplan.2019.04.024)
- Pretty, J., Rogerson, M., & Barton, J. (2017). Green mind theory: How brain-body-behaviour links into natural and social environments for healthy habits. *International Journal of Environmental Research and Public Health*, 14(7), 706.  
DOI: [10.3390/ijerph14070706](https://doi.org/10.3390/ijerph14070706)
- Qiu, L., Lindberg, S., & Nielsen, A. B. (2013). Is biodiversity attractive?—On-site perception of recreational and biodiversity values in urban green space. *Landscape and Urban Planning*, 119, 136–146.  
DOI: [10.1016/j.landurbplan.2013.07.007](https://doi.org/10.1016/j.landurbplan.2013.07.007)

- Reagan, A. J., Danforth, C. M., Tivnan, B., Williams, J. R., & Dodds, P. S. (2017). Sentiment analysis methods for understanding large-scale texts: a case for using continuum-scored words and word shift graphs. *EPJ Data Science*, 6(1), 28.  
DOI: [10.1140/epjds/s13688-017-0121-9](https://doi.org/10.1140/epjds/s13688-017-0121-9)
- Rigolon, A. (2016). A complex landscape of inequity in access to urban parks: A literature review. *Landscape and Urban Planning*, 153, 160–169.  
DOI: [10.1016/j.landurbplan.2016.05.017](https://doi.org/10.1016/j.landurbplan.2016.05.017)
- Rigolon, A., Browning, M., & Jennings, V. (2018). Inequities in the quality of urban park systems: An environmental justice investigation of cities in the united states. *Landscape and Urban Planning*, 178, 156–169.  
DOI: [10.1016/j.landurbplan.2018.05.026](https://doi.org/10.1016/j.landurbplan.2018.05.026)
- Roberts, H. [Hannah], Kellar, I., Conner, M., Gidlow, C., Kelly, B., Nieuwenhuijsen, M., & McEachan, R. (2019). Associations between park features, park satisfaction and park use in a multi-ethnic deprived urban area. *Urban Forestry & Urban Greening*, 46, 126485.  
DOI: [10.1016/j.ufug.2019.126485](https://doi.org/10.1016/j.ufug.2019.126485)
- Roberts, H. [Helen], Resch, B., Sadler, J., Chapman, L., Petutschnig, A., & Zimmer, S. (2018). Investigating the emotional responses of individuals to urban green space using twitter data: A critical comparison of three different methods of sentiment analysis. *Urban Planning*, 3(1), 21.  
DOI: [10.17645/up.v3i1.1231](https://doi.org/10.17645/up.v3i1.1231)
- Sato, M., Inoue, Y., Du, J., & Funk, D. C. (2019). Access to parks and recreational facilities, physical activity, and health care costs for older adults: Evidence from us counties. *Journal of Leisure Research*, 50(3), 220–238.  
DOI: [10.1080/00222216.2019.1583048](https://doi.org/10.1080/00222216.2019.1583048)
- Saw, L. E., Lim, F. K. S., & Carrasco, L. R. (2015). The relationship between natural park usage and happiness does not hold in a tropical city-state. *PLOS ONE*, 10(7), e0133781.  
DOI: [10.1371/journal.pone.0133781](https://doi.org/10.1371/journal.pone.0133781)
- Schwartz, A. J., Dodds, P. S., O’Neil-Dunne, J. P., Danforth, C. M., & Ricketts, T. H. (2019). Visitors to urban greenspace have higher sentiment and lower negativity on twitter. *People and Nature*, 1(4), 476–485.  
DOI: [10.1002/pan.3.10045](https://doi.org/10.1002/pan.3.10045)
- Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and statistical modeling with python. In *9th python in science conference*.
- Shanahan, D. F., Fuller, R. A., Bush, R., Lin, B. B., & Gaston, K. J. (n.d.). The health benefits of urban nature: how much do we need? *BioScience*, (5), 476–

485.

DOI: [10.1093/biosci/biv032](https://doi.org/10.1093/biosci/biv032)

Soga, M., & Gaston, K. J. (2016). Extinction of experience: The loss of human–nature interactions. *Frontiers in Ecology and the Environment*, 14(2), 94–101.

DOI: [10.1002/fee.1225](https://doi.org/10.1002/fee.1225)

South, E. C., Hohl, B. C., Kondo, M. C., MacDonald, J. M., & Branas, C. C. (2018). Effect of greening vacant land on mental health of community-dwelling adults: A cluster randomized trial. *JAMA network open*, 1(3), e180298–e180298.

DOI: [10.1001/jamanetworkopen.2018.0298](https://doi.org/10.1001/jamanetworkopen.2018.0298)

Stommel, M., & Schoenborn, C. A. (2009). Accuracy and usefulness of bmi measures based on self-reported weight and height: Findings from the nhanes & nhis 2001–2006. *BMC public health*, 9(1), 421.

DOI: [10.1186/1471-2458-9-421](https://doi.org/10.1186/1471-2458-9-421)

Taylor, L., Hahs, A. K., & Hochuli, D. F. (2018). Wellbeing and urban living: Nurtured by nature. *Urban Ecosystems*, 21(1), 197–208.

DOI: [10.1007/s11252-017-0702-1](https://doi.org/10.1007/s11252-017-0702-1)

The Trust for Public Land. (2019). Park serve. data retrieved from TPL, <https://www.tpl.org/parkserve/downloads>.

Tsai, W.-L., Davis, A. J., & Jackson, L. E. (2019). Associations between types of greenery along neighborhood roads and weight status in different climates. *Urban forestry & urban greening*, 41, 104–117.

DOI: [10.1016/j.ufug.2019.03.011](https://doi.org/10.1016/j.ufug.2019.03.011)

Tsai, W.-L., McHale, M. R., Jennings, V., Marquet, O., Hipp, J. A., Leung, Y.-F., & Floyd, M. F. (2018). Relationships between characteristics of urban green land cover and mental health in us metropolitan areas. *International journal of environmental research and public health*, 15(2), 340.

DOI: [10.3390/ijerph15020340](https://doi.org/10.3390/ijerph15020340)

Twitter Streaming API. (2016). <https://developer.twitter.com>.

U.S. Census Bureau. (2012). U.S. Census Populated Places. data retrieved from ESRI, <https://hub.arcgis.com/datasets/esri::usa-census-populated-places>.

U.S. Census Bureau. (2017). American Community Survey (ACS): 5-Year Estimates 2013–2017. <https://www.census.gov/programs-surveys/acs>.

Ulrich, R. S., Simons, R. F., Losito, B. D., Fiorito, E., Miles, M. A., & Zelson, M. (1991). Stress recovery during exposure to natural and urban environments. *Journal of Environmental Psychology*, 11(3), 201–230.

DOI: [10.1016/S0272-4944\(05\)80184-7](https://doi.org/10.1016/S0272-4944(05)80184-7)

- United Nations. (2014). World Urbanization Prospects: The 2014 Revision-Highlights. <https://www.un.org/en/development>. United Nations.
- Van den Berg, A. E. (2017). From green space to green prescriptions: Challenges and opportunities for research and practice. *Frontiers in psychology*, 8, 268.  
DOI: [10.3389/fpsyg.2017.00268](https://doi.org/10.3389/fpsyg.2017.00268)
- Van den Berg, M., Wendel-Vos, W., van Poppel, M., Kemper, H., van Mechelen, W., & Maas, J. (2015). Health benefits of green spaces in the living environment: A systematic review of epidemiological studies. *Urban Forestry & Urban Greening*, 14(4), 806–816.  
DOI: [10.1016/j.ufug.2015.07.008](https://doi.org/10.1016/j.ufug.2015.07.008)
- van den Berg, M., van Poppel, M., van Kamp, I., Andrusaityte, S., Balseviciene, B., Cirach, M., . . . Maas, J. (2016). Visiting green space is associated with mental health and vitality: A cross-sectional study in four european cities. *Health & Place*, 38, 8–15.  
DOI: [10.1016/j.healthplace.2016.01.003](https://doi.org/10.1016/j.healthplace.2016.01.003)
- van den Bosch, M., & Nieuwenhuijsen, M. (2017). No time to lose—green the cities now. *Environment International*, 99, 343–350.  
DOI: [10.1016/j.envint.2016.11.025](https://doi.org/10.1016/j.envint.2016.11.025)
- van den Bosch, M., & Sang, Å. O. (2017). Urban natural environments as nature-based solutions for improved public health—a systematic review of reviews. *Environmental research*, 158, 373–384.  
DOI: [10.1016/j.envres.2017.05.040](https://doi.org/10.1016/j.envres.2017.05.040)
- Wang, H., & Tassinary, L. G. (2019). Effects of greenspace morphology on mortality at the neighbourhood level: A cross-sectional ecological study. *The Lancet Planetary Health*, 3(11), e460–e468.  
DOI: [10.1016/S2542-5196\(19\)30217-7](https://doi.org/10.1016/S2542-5196(19)30217-7)
- Wang, Y., Kotze, D. J., Vierikko, K., & Niemelä, J. (2019). What makes urban greenspace unique—relationships between citizens’ perceptions on unique urban nature, biodiversity and environmental factors. *Urban Forestry & Urban Greening*, 42, 1–9.  
DOI: [10.1016/j.ufug.2019.04.005](https://doi.org/10.1016/j.ufug.2019.04.005)
- Ward Thompson, C., Aspinall, P., Roe, J., Robertson, L., & Miller, D. (2016). Mitigating stress and supporting health in deprived urban communities: The importance of green space and the social environment. *International Journal of Environmental Research and Public Health*, 13(4), 440.  
DOI: [10.3390/ijerph13040440](https://doi.org/10.3390/ijerph13040440)
- Water Bodies. (2019). <https://data.sfgov.org/>. San Francisco Department of Public Works, San Francisco, CA.

- Waters, H., & Graf, M. (2018). The Costs of Chronic Disease in the US. <https://milkeninstitute.org/reports/costs-chronic-disease-us>. Milken Institute.
- Wen, M., Zhang, X., Harris, C. D., Holt, J. B., & Croft, J. B. (2013). Spatial disparities in the distribution of parks and green spaces in the usa. *Annals of Behavioral Medicine*, 45(suppl\_1), S18–S27.  
DOI: [10.1007/s12160-012-9426-x](https://doi.org/10.1007/s12160-012-9426-x)
- West, S. T., Shores, K. A., & Mudd, L. M. (2012). Association of available parkland, physical activity, and overweight in america's largest cities. *Journal of Public Health Management and Practice*, 18(5), 423–430.  
DOI: [10.1097/PHH.0b013e318238ea27](https://doi.org/10.1097/PHH.0b013e318238ea27)
- Wheeler, B. W., Lovell, R., Higgins, S. L., White, M. P., Alcock, I., Osborne, N. J., ... Depledge, M. H. (2015). Beyond greenspace: an ecological study of population general health and indicators of natural environment type and quality. *International Journal of Health Geographics*, 14(1), 17.  
DOI: [10.1186/s12942-015-0009-5](https://doi.org/10.1186/s12942-015-0009-5)
- White, M. P., Alcock, I., Grellier, J., Wheeler, B. W., Hartig, T., Warber, S. L., ... Fleming, L. E. (2019). Spending at least 120 minutes a week in nature is associated with good health and wellbeing. *Scientific Reports*, 9(1), 7730.  
DOI: [10.1038/s41598-019-44097-3](https://doi.org/10.1038/s41598-019-44097-3)
- Wilkie, S., Thompson, E., Cranner, P., & Ginty, K. (2020). Attention restoration theory as a framework for analysis of tweets about urban green space: A case study. *Landscape Research*, 1–12.  
DOI: [10.1080/01426397.2020.1738363](https://doi.org/10.1080/01426397.2020.1738363)
- Zhang, J. W., Piff, P. K., Iyer, R., Koleva, S., & Keltner, D. (2014). An occasion for unselfing: Beautiful nature leads to prosociality. *Journal of Environmental Psychology*, 37, 61–72.  
DOI: [10.1016/j.jenvp.2013.11.008](https://doi.org/10.1016/j.jenvp.2013.11.008)
- Zhang, L., Tan, P. Y., & Diehl, J. A. (2017). A conceptual framework for studying urban green spaces effects on health. *Journal of Urban Ecology*, 3(1), jux015.  
DOI: [10.1093/jue/jux015](https://doi.org/10.1093/jue/jux015)
- Zhang, S., & Zhou, W. (2018). Recreational visits to urban parks and factors affecting park visits: Evidence from geotagged social media data. *Landscape and Urban Planning*, 180, 27–35.  
DOI: [10.1016/j.landurbplan.2018.08.004](https://doi.org/10.1016/j.landurbplan.2018.08.004)
- Zhang, X., Holt, J. B., Lu, H., Wheaton, A. G., Ford, E. S., Greenlund, K. J., & Croft, J. B. (2014). Multilevel regression and poststratification for small-area estimation of population health outcomes: A case study of chronic obstructive pulmonary disease prevalence using the behavioral risk factor surveillance

system. *American journal of epidemiology*, 179(8), 1025–1033.

DOI: [10.1093/aje/kwu018](https://doi.org/10.1093/aje/kwu018)

Zhang, X., Holt, J. B., Yun, S., Lu, H., Greenlund, K. J., & Croft, J. B. (2015). Validation of multilevel regression and poststratification methodology for small area estimation of health indicators from the behavioral risk factor surveillance system. *American journal of epidemiology*, 182(2), 127–137.

DOI: [10.1093/aje/kwv002](https://doi.org/10.1093/aje/kwv002)

# Chapter 7: Appendix

## 7.1 Chapter 3 Appendix

### 7.1.1 Twitter API

Twitter’s ‘spritzer’ streaming API offers a random selection of up to 1% of all messages, with specific linguistic or spatial filters enabling a higher percentage. For the present study, we collected messages tagged with GPS coordinates during the years 2012–2015. During this period, geolocated messages comprised roughly 1% of all messages. As a result, filtering on GPS enabled us to collect nearly 100% of all such messages.

### 7.1.2 Stopwords

As is common in natural language processing, we define ‘stop words’ as individual words that we mask from sentiment analysis. These are words that we identify as frequent in our tweets, but that contribute neutral or context-dependent sentiment. We do not include the word *park* in our analysis. We removed the words *closed*, *traffic*, and *accident* because they frequently appeared in geo-located tweets from automated traffic posts. We removed words found in the names of the parks (e.g., *golden* and *gate*). Several cities had increased frequencies for the positive words *art*, *museums*, *gardens*, and *zoos* in their parks. Even though these words were not in the official park names, we removed them from our analysis. Several parks had the positive words *music* and *festival* appear frequently, so we removed these two words. For each city, we identified a list of stop words to remove by manually checking the 10 most influential words contributing to the difference between in-park and control tweets. Finally, we removed words that referred to a specific location (e.g., *beach*) or were being used in a significantly different way than they were originally rated for happiness (e.g. *ma* as shorthand for Massachusetts rather than mother) were removed (See Table 7.1).

Overall, the majority of words we masked were positive, with average happiness scores greater than 6 as seen in Fig. 7.1. As a result, we expect that the happiness benefit reported in our results is a lower bound.

### 7.1.3 Hashtags

Tweets with any of the following hashtags were removed from our study sample: #jobs, #job, #getalljobs, #hiring, #tweetmyjobs, #careerarc, #hospitality, #healthcare, #nursing, #marketing, #sales, #clerical, #it.



*Table 7.1: Stop words selected for individual cities based on frequency analysis and contextual meaning.*

<u>City</u>	<u>Stop Words</u>
San Francisco	young, flowers
Phoenix	hospital
Jacksonville	science
Austin	limits
San Diego	sea
Washington	war, bill, united, health
Seattle	health, surgery, emergency
Chicago	riot
Houston	hospital, delay, stop, science
Cleveland	beach, island
Boston	ma, partners
New York	natural
San Antonio	cafe
Dallas	health
Philadelphia	independence
Los Angeles	science
San Jose	christmas, raging
Denver	nature, science, international
Memphis	steal, sugar
Charlotte	shot, young
Indianapolis	health
Columbus	roses

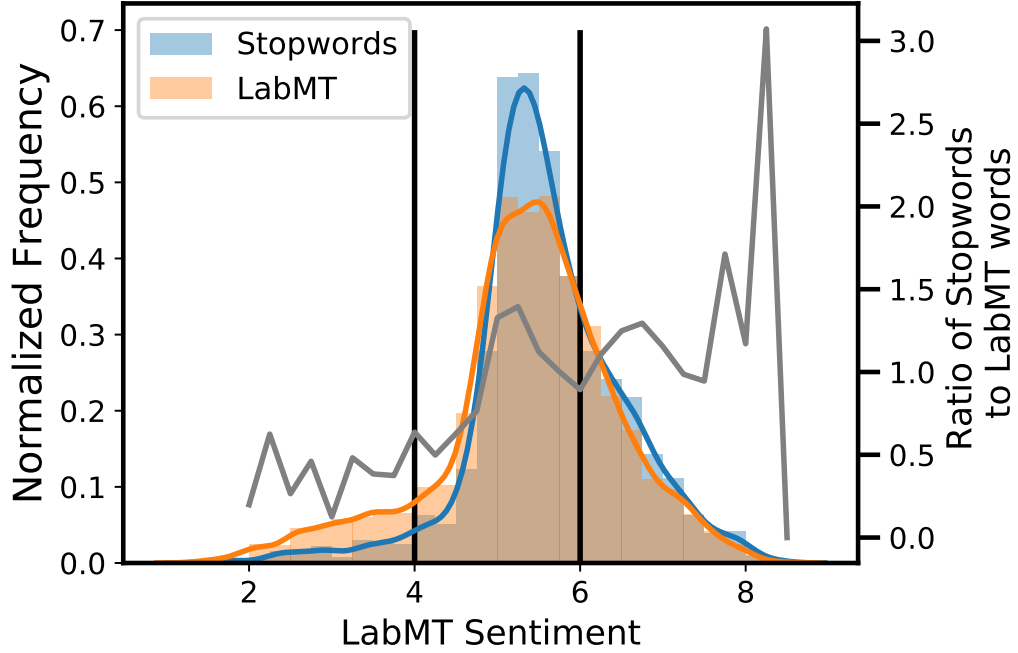


Figure 7.1: Normalized histogram of LabMT words and stop words taken out of the analysis due to being in a park name. Our analysis is conservative as the ratio is higher for positive words ( $> 6$ ) compared to negative words ( $< 4$ ). Words between 4 and 6 are not included in our analysis.

#### 7.1.4 Happiness Benefit: User Control

In addition to the proximate time control described in the Methods section, we employed a secondary control to investigate the happiness benefit methodology. In this method, we selected a ‘user control’ tweet: a random message from the same user posted out-of-park. If an account’s message history consisted entirely of in-park tweets, the account was removed from the sample as they were likely a tourist or business located adjacent to the park. The user control allows us to estimate a happiness benefit for the users during their park visits compared to tweets when they were not in the parks. We performed the same happiness benefit calculation for each of the 25 cities and include those results in Figure 7.2. For our ‘user control’ group, the mean happiness benefit for the cities in our sample ranged from  $-0.02$  to  $.05$  (Fig. 7.2). We also plot the mean happiness benefit against park spending for capita and Park Score® in Fig. 7.3. The overall benefit reduction observed for the User Control, when compared with the time control, suggests that individuals who tweet from within parks generally use happier words than individuals who do not visit parks.

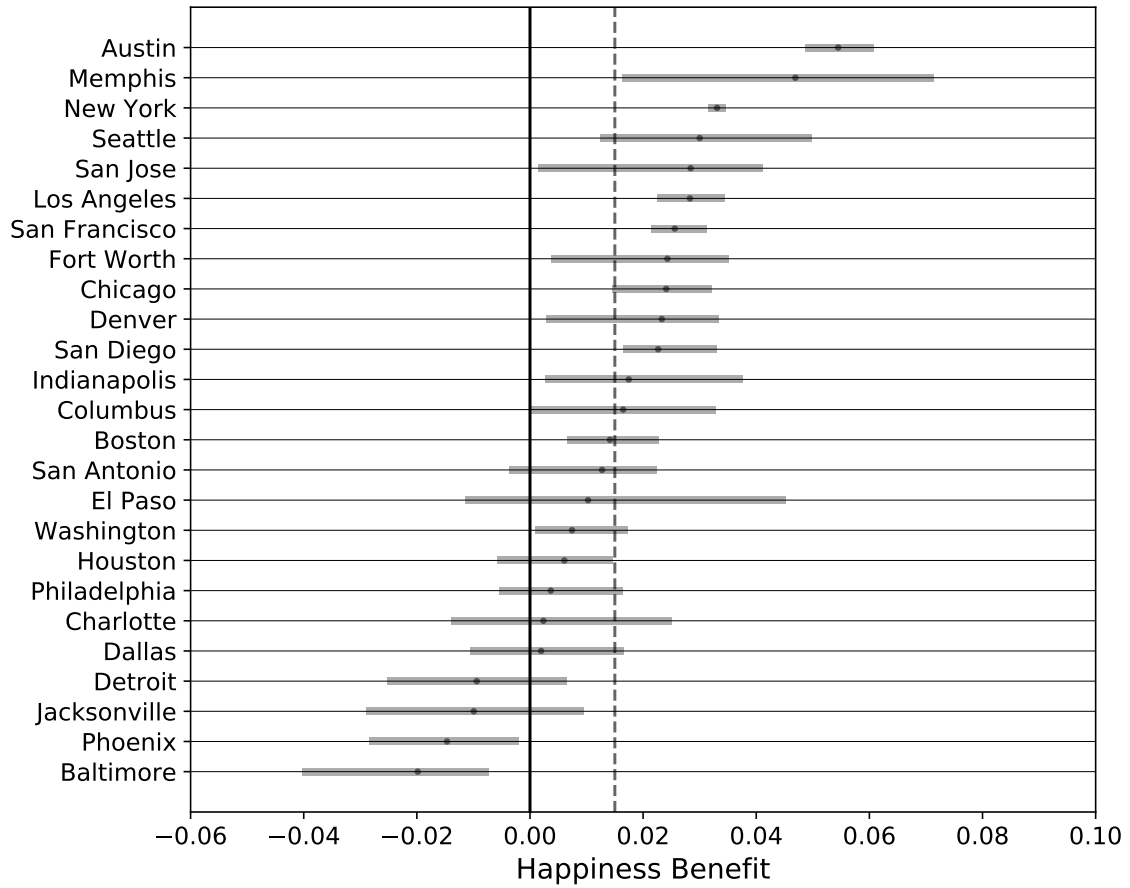


Figure 7.2: Happiness benefit by city. We derive each city's full range of values from 10 bootstrap runs, for which we randomly selected 80% of tweets. Darker dots represent mean value from bootstrap runs. For each city, the control group consists of 1 random, non-park tweet from each user paired with an in-park tweet.

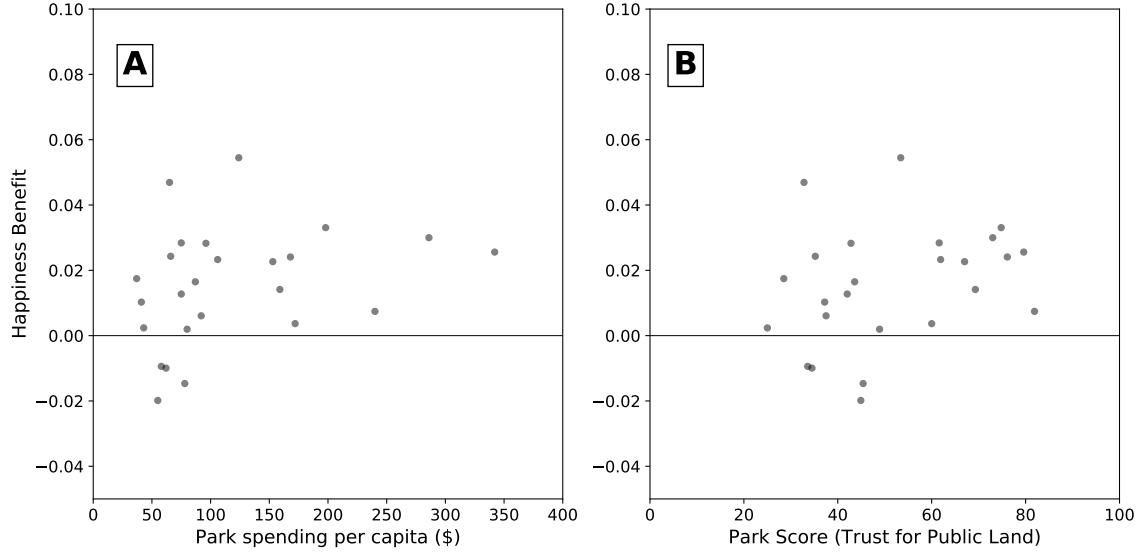


Figure 7.3: A. The left panel shows park spending per capita vs mean happiness benefit by city. Park spending per capita is from Trust for Public Land (TPL) data. B. The right panel shows ParkScore® vs mean happiness. The TPL calculates ParkScore® annually from measures of park acreage, access, investment, and amenities, and is scaled to a maximum score of 100.

### 7.1.5 Temporal Analysis by hour of day

We estimated the happiness benefit by hour of day across all cities (Fig. 7.4). While 8:00 and 9:00AM are slightly lower, the rest of the day's happiness benefit ranges overlap, showing that our other results are not biased by certain hours of the day (e.g., leaving the office).

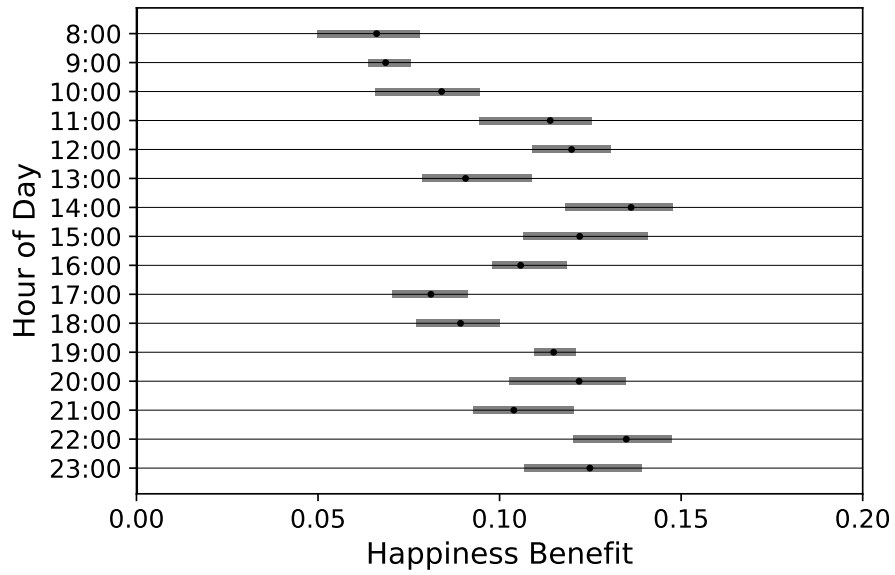


Figure 7.4: Change in happiness benefit by hour of day. The range is the full range of happiness benefit estimates from 10 runs, sampling 80% of tweets. 1,000 random in-park tweets were pooled in each group from each city. Control tweets were selected as tweets most temporally proximate to the in-park tweet from the same city.

## **7.2 Chapter 4 Appendix**

### **7.2.1 500 Cities Data**

The CDC estimates three categories of data for the 500 Cities project: health outcomes, prevention, and unhealthy behaviors. The CDC classifies obesity under unhealthy behaviors, as it is a risk factor for several chronic diseases including heart disease and stroke (CDC, 2018). Frequent Mental Distress (FMD) is categorized under health outcomes.

### **7.2.2 500 Cities Validation**

The CDC 500 Cities health estimates have been validated using observed prevalence data from other surveys. Direct surveys from Boston were compared with estimates based on BRFSS data and concluded that estimates were both valid and useful for characterizing geographic variation in health outcomes (Zhang et al., 2015). In another study, correlation coefficients between CDC estimates and direct surveys for Chronic Obstructive Pulmonary Disease Prevalence (COPD) ranged from .88 to .95 at the County Level. Contextual effects - modeled as state and county level random effects - were significant, and not explainable by demography alone (Zhang et al., 2014). According to the CDC, their model estimates do not account for local policy or program intervention effects which may introduce bias (CDC, 2018). An earlier study found that including contextual effects related to the physical environment could improve modeled results for small-area estimates of obesity. However, authors called for careful study before including them due to potential non-linear effects (Li et al., 2009).

### **7.2.3 Small-Area Estimation Methodology**

The BRFSS conducts annual telephone surveys across the entire US on health conditions and health-related behaviors. To estimate the 500 Cities prevalence data, the CDC models health outcomes at different geographic levels by linking the BRFSS survey results with population level demographic and socioeconomic data (poverty status, education, and race/ethnicity) using a method called small-area estimation (Li et al., 2009; Zhang et al., 2014). In other words, the 500 Cities data are expected estimates of a health outcome conditional on the socioeconomic and demographic makeup of that city along with unmeasured contextual factors at the county and state level.